GROWING UP IN SCOTLAND: CHANGES IN LANGUAGE ABILITY OVER THE PRIMARY SCHOOL YEARS

Growing Up in Scotland

GROWING UP IN SCOTLAND: CHANGES IN LANGUAGE ABILITY OVER THE PRIMARY SCHOOL YEARS

Authors: Line Knudsen[1], Eilidh Currie[1], Paul Bradshaw[1], James Law[2], Rachael Wood[3]

[1] ScotCen Social Research

[2] Department of Speech and Language, University of Newcastle

[3] Information Services Division Scotland and University of Edinburgh

ACKNOWLEDGEMENTS

First and foremost, the authors would like to thank all the families who have given up their time to take part in the study and have supported it continuously sweep by sweep.

The Growing Up in Scotland study is very much a collaborative venture. We owe a huge debt of gratitude to our colleagues in NatCen's operations and computing departments and to our team of interviewers and supervisors for conducting the interviews with such dedication and professionalism.

Responsibility for the opinions expressed in this report, and for all interpretation of the data, lies solely with the authors.

GROWING UP IN SCOTLAND

CHANGES IN LANGUAGE ABILITY OVER THE PRIMARY SCHOOL YEARS

CONTENTS

EXECUTIVE SUMMARY

Overview and key aims of the report

This report draws on measures of expressive language ability which were obtained for children participating in the Growing Up in Scotland study (GUS) at the time they were about to or had recently entered primary school (in 2009/10) and again when they were in Primary 6 (in 2014/15).

Building on what is already known about differences in language ability and what might influence this up until entry to primary school, the report explores first whether there remains a 'gap' in expressive language ability between more and less advantaged children towards the end of the primary school period. It also considers whether the gap appears to have changed since the children started primary school. The report then identifies characteristics, circumstances and experiences present over the primary school years which appear to help or hinder children's expressive language development, relative to their peers. In doing so, the report helps us understand more about what might help children improve during this period. The findings may therefore help policy makers and others target their efforts to reduce the attainment gap, as well as pointing to avenues for further research.

Is there still a gap in expressive language ability between more and less advantaged children towards the end of primary school?

The findings in this report indicate that the gap between more and less advantaged children persists and is evident as children reach the last years of primary school. This is the case irrespective of whether the gap is measured in relation to differences by family income, the level of area deprivation or parental level of education.

Furthermore, the findings indicate that, if anything, this gap appears to have widened rather than narrowed since the children entered primary school. However, the data and analytical approach mean that we are not able to estimate by what margin the gap has widened.

Notably, despite showing significant inequalities in average language ability between different groups of children – such as those in the highest and those in the lowest income households – the report also found evidence of significant variation in ability within these groups. For example, many children in lower income households had relatively high language ability whilst many in higher income households had relatively low ability. This suggests that whilst being from a disadvantaged social background increases the risk of poorer language skills, it does not necessarily equate to poorer language skills for *all* children from disadvantaged backgrounds. Similarly, and importantly, coming from a more advantaged background does not guarantee more advanced language development.

What might help (and what seems to hinder) improvement in language skills over the primary school period?

The analysis only explains a limited amount of why some children show greater improvement in their language ability of the primary school period. Nevertheless, some factors which do appear to be associated with improvement are worth noting. First, the findings suggest that children with above average levels of social, emotional or behavioural difficulties tend to show a decline in their language ability relative to those children without social, emotional or behavioural difficulties, even when taking into account several other known differences between the children. This highlights the importance of policies and initiatives aimed at supporting children's educational attainment taking into account other aspects of the child's development – including their mental health and wellbeing. It also stresses the need to ensure that children with additional support needs associated with social and behavioural development are fully supported throughout primary school.

Second, the report showed a positive relationship between frequent home reading when the child was approaching 8 years (when most children were in Primary 4) and a higher level of improvement in expressive language ability relative to their peers over the primary school period, including when other known differences such as parental education were taken into account. This may reflect that children who experience an improvement in their language ability develop (more of) an interest in reading. However, it may also be an indication that frequent home learning activities such as reading continue to have benefits for children's language development beyond the early years, thus lending support to initiatives aimed at encouraging parents of school-aged children to engage their child(ren) in these activities.

Finally, the findings indicate that over the primary school period the expressive language skills of children living in small towns and rural areas improve at a higher rate than those of children living in urban areas, also when accounting for differences in other characteristics and circumstances, including differences in their parents' level of education. Whilst the report does not identify *why* such differences occur – for example, which systematic differences in growing up in small town and rural areas are particularly important for children's expressive language development – this would be a useful avenue for further research to explore.

1 INTRODUCTION

1.1. Background and report overview

Improving educational attainment and closing the poverty-related attainment gap has been high on the agenda for policymakers in Scotland for several years. It is also a prominent priority for campaign groups and charities such as Save the Children. This is supported by existing evidence demonstrating that children from poorer families tend to have poorer educational outcomes than those from more affluent families (Sosu and Ellis, 2014).

There is evidence to suggest that one of the key factors driving this attainment gap is the high prevalence of early difficulties in language ability among disadvantaged children (Law et al., 2017). Language ability during the formative years has long been recognised as important for later attainment and outcomes. A considerable body of research has demonstrated that poor early language ability is associated with low educational attainment, in turn affecting individuals' employment prospects and health (Howieson and Iannelli, 2008; Ritchie et al., 2015).

Using Growing Up in Scotland (GUS) data, Bradshaw (2011) examined the gap in language ability among pre-school children in Scotland and identified some of the factors most strongly associated with relative improvement between the ages of 3 and 5. More recent data from GUS, collected at the time the children involved in the study were in Primary 6, offers an opportunity to further explore changes in children's language ability across the primary school years.

This report draws on measures of expressive language ability obtained for the GUS children at the time they were about to or had recently entered primary school (in 2009/10) and at the time they were in Primary 6 (in 2014/15). Building on what is already known about differences in language ability and factors influencing language ability up until entry to primary school, this report examines the gap in expressive vocabulary ability towards the end of primary school period, and identifies factors present over the primary school years which appear to help or hinder children's language development over this period, relative to their peers. Given the link between language ability and attainment, in doing so, the research adds to the evidence base on how to improve attainment for children in Scotland; – understanding more about what might help children improve can help policy makers and others target their efforts and can also point towards avenues for further research.

1.2. The poverty-related attainment gap in Scotland: what do we know?

Previous research from GUS identified a developmental gap among children with different background characteristics even before they had started school (Bradshaw, 2011). The report examined changes in the cognitive ability of children aged 3 and 5

from different social backgrounds. It showed that, at both ages, children from more advantaged households significantly outperformed those from less advantaged households on measures of expressive vocabulary and problem solving, with differences in children's cognitive ability according to their parents' level of education, income and social class. Regarding expressive language, differences in ability by parental education were particularly prominent, with those whose parent(s) had higher qualification levels demonstrating better vocabulary than those whose parent(s) had lower levels or no qualifications. Substantive differences in knowledge of vocabulary were also evident by income and social class. Overall, this report demonstrated that the attainment gap among children in Scotland is already evident by the age of 3 and appears to widen in certain domains of learning by age 5.

There is also evidence to suggest that the gap persists across the school years. For example, the most recent results from the Scottish Survey of Literacy and Numeracy (SSLN)[1] published in 2017, showed that on all four literacy indicators - reading, writing, listening and talking - pupils from the least deprived areas outperformed pupils from the most deprived areas, at all stages (Scottish Government, 2017a). For example, the proportion of Primary 4 children who were assessed as doing well or very well in reading was 67% in the most deprived areas compared with 85% in the least deprived areas. Similar patterns were evident for writing, listening and talking.

Existing research has also considered the attainment of school leavers, based on data from the pupil census and the Scottish Qualifications Authority. The most recent data show a gap in attainment between leavers from the most and least deprived areas in Scotland (Scottish Government, 2016). In 2015/16, 99% of leavers from the 20% least deprived areas (using the Scottish Index of Multiple Deprivation (SIMD)) obtained a qualification at SCQF level 4[2] or above compared with 93% of leavers from the 20% most deprived areas. The gap is larger at SCQF level 6 or above[3], with 81% of leavers from the 20% least deprived areas obtaining a qualification at this level or above, compared with 43% from the 20% most deprived areas.

Thus, existing research suggests that across a range of different measures and at different stages of childhood and adolescence, attainment and ability in Scotland is stratified by deprivation, with children from less advantaged backgrounds achieving poorer educational outcomes than those from more advantaged backgrounds. This report builds on and adds to this evidence through examining children's expressive language development over the primary school years.

[1] More detailed information about the SSLN can be found here: http://www.gov.scot/Topics/Statistics/Browse/School-Education/SSLN

[2] Equivalent to National 4 / Intermediate 1 / Standard Grade General.

[3] Equivalent to a Higher or above.

1.3. Closing the poverty-related attainment gap in Scotland: the policy context

The Scottish Government has introduced a range of policies, strategies and frameworks designed to close the poverty-related attainment gap. Most significant of these is the Scottish Attainment Challenge which was launched in 2015. The Scottish Attainment Challenge is a national initiative which aims to reduce inequity by improving educational outcomes for children living in Scotland's most disadvantaged communities. It has a focus on supporting schools and local authorities to improve outcomes in literacy and numeracy, as well as health and wellbeing. There is a specific emphasis on those living in the 20% most deprived areas in Scotland. The Scottish Attainment Challenge comprises a range of initiatives, including extra money for schools in deprived areas and councils, an Attainment Advisor in every local authority to help schools and teachers and an online 'hub' to help educationalists find examples of good practice.

A key element of the Scottish Attainment Challenge is the £750 million Attainment Scotland Fund, a targeted initiative focused on supporting pupils in the local authorities of Scotland with the highest concentrations of deprivation. The nine 'Challenge Authorities' are Glasgow, Dundee, Inverclyde, West Dunbartonshire, North Ayrshire, Clackmannanshire, North Lanarkshire, East Ayrshire and Renfrewshire. Pupil Equity Funding is also provided through the Attainment Scotland Fund and allocated directly to schools based on the estimated number of children and young people in P1-S3 registered for free school meals under the national eligibility criteria[4]. The central aim of the Scottish Attainment Challenge is to achieve long term educational improvement and opportunities for children living in areas of multiple deprivation.

In addition, the National Improvement Framework for Scottish education (Scottish Government, 2017b) is designed to secure educational improvement in Scotland. Key aims of this policy include improving attainment in literacy and numeracy and closing the gap between the most and least disadvantaged children, as well as improving children and young people's health, wellbeing and employability skills. The National Improvement Framework sets out six key drivers for improvement. These include school leadership; teacher professionalism; parental engagement; assessment of children's progress; school improvement and performance information. Parental engagement is highlighted as a key factor to help children achieve the highest standards whilst reducing inequity and closing the attainment gap, with evidence from the annual statutory review of the National Improvement Framework in 2016 showing that family learning helps close the attainment gap through breaking the intergenerational cycles of deprivation and low attainment.

The Scottish Attainment Challenge and the National Improvement Framework are underpinned by a broader range of Scottish Government initiatives and programmes

[4] http://www.gov.scot/Topics/Education/Schools/HLivi/schoolmeals/FreeSchoolMeals

which, though not specifically focused on weakening the link between poverty and low educational attainment, could enable educational establishments to address the impact of disadvantage on educational attainment.

These include Curriculum for Excellence (CfE). Introduced in 2006, CfE is Scotland's curriculum for learners aged 3-18 which sets out the aims, principles and approaches that should underpin learning for those aged 3 to 18 years in Scotland. CfE has two phases: the broad general education (from the early years to the end of S3) and the senior phase (S4 to S6).

CfE offers several important themes to enhance the delivery of education to disadvantaged groups (Scottish Government, 2008). For example, literacy, numeracy and health and wellbeing are recognised as being particularly important and the responsibility of all staff. In addition, CfE promotes flexibility, personalisation and choice, and challenges schools and their partners to support children to become 'successful learners', 'confident individuals', 'responsible citizens' and 'effective contributors'.

Sosu and Ellis (2014) argue that if CfE is tailored to meet the educational needs of children from deprived households, it could be a powerful force for closing the poverty-related attainment gap in Scotland. However, the dual aim of closing the gap and at the same time raising the bar for all children has led to concerns that privileged students, parents, schools and communities will be more likely to make progress, due to the considerable discretion which can be exercised in the implementation of CfE (OECD, 2015).

Getting it right for every child (GIRFEC) - the national approach to improving the wellbeing of children and young people in Scotland (Scottish Government, 2018) - also addresses issues of disadvantage and educational attainment. GIRFEC is designed to ensure that all children and young people are offered the help that may support them to be successful in life, including at school. The framework focuses attention on how schools, working with families and their partners, might better meet the needs of all learners, including those from socio-economically disadvantaged backgrounds. The wellbeing indicators (Safe, Healthy, Active, Nurtured, Achieving, Respected, Responsible and Included) have encouraged a focus on disadvantaged groups. GIRFEC also promotes support for individual children and young people through a staged intervention mechanism, which provides a framework for additional targeted support to meet their wellbeing needs. Although not specifically designed to close the poverty-related attainment gap, consideration of a child's or young person's wellbeing includes taking account of environmental circumstances like living in poverty and it has been argued that these measures have the potential to prompt schools and others to address the educational disparities that arise from economic disadvantage (Sosu and Ellis, 2014).

1.4. About the Growing Up in Scotland study (GUS)

GUS is a longitudinal research study which tracks the lives of thousands of children and their families in Scotland from the early years, through childhood and beyond. The main aim of the study is to provide new information to support policy-making in Scotland, but it is also intended to provide a resource for practitioners, academics, the voluntary sector and parents.

To date, the study has collected information about three nationally representative cohorts of children: a child cohort and two birth cohorts. Altogether, information has been collected on around 14,000 children and families in Scotland.

This report draws on data collected at the time children in the first GUS birth cohort were about to or had recently entered primary school (2009/10) and at the time they were in Primary 6 (2014/15). More detailed information about the data is provided in section 2.1.

1.5. Research questions

Taking a similar approach to the previous GUS report examining changes in cognitive ability in the pre-school years (Bradshaw, 2011), this report focuses on changes in language ability over the primary school years. It addresses the following questions:

1. Does the gap in expressive language ability between children from advantaged and disadvantaged backgrounds change over the primary school years?
2. What circumstances and experiences are associated with a relative change in expressive language ability over the primary school years?
3. Do the factors associated with a relative change in ability vary according to social background?

The gap in expressive language ability (according to income, area deprivation and parental education) among children in Primary 6, towards the end of primary school, is considered in chapter 3.

Chapter 4 explores what experiences and circumstances are associated with a relative improvement or decline in expressive language ability over the primary school period, and whether this differs according to parental education.

Finally, chapter 5 draws together key findings from the previous chapters and suggests what implications they have for policy makers and others seeking to improve language development for children in Scotland.

2 METHODS

2.1. Sample and data overview

The analysis presented in this report uses data from the first GUS birth cohort (Birth Cohort 1 or 'BC1'). BC1 is comprised of a nationally representative sample of 5217 children living in Scotland when they were 10 months old who were born between June 2004 and May 2005.

Starting in 2005/06, data were collected annually from when the children were aged 10 months until they were just under 6 years old, and then biennially at age 7-8 and when the children were in Primary 6 (age 10-11). At the time of writing (2018), the ninth sweep of face-to-face data collection with this cohort has finished. At this ninth sweep the cohort children were in their first year of secondary school (age 12-13). This report draws primarily on data collected, firstly, at the time the children were aged just under 5 when most children were in Primary 1 or about to enter primary school (in 2009/10), and secondly, at the time they were in Primary 6 and aged 10-11 (in 2014/15), although for a number of the factors examined in chapter 4 it also draws on data collected out with these two time points (see Table 2-1 for details). Because the cohort is comprised of a nationally representative sample of children the results should be understood to represent all children of the respective age living in Scotland at the time point in question who were also living in Scotland when they were 10 months old. For example, the results presented for the GUS children at the time they were in Primary 6 are roughly representative of all children in Scotland who attended Primary 6 in 2014/15[5].

The main data collection on GUS takes place through annual or biennial 'sweeps' of face-to-face interviews with children and parents in their homes. This report draws on data collected from several sources. First, it draws on data collected from the cohort child's main carer at various age points. Second, it draws on objective measures of the child's vocabulary at the time most children were in or about to enter Primary 1 and when they were in Primary 6 (see further details in section 2.2). Third, it draws on data collected from the children themselves when they were aged 8, around the time most children were in Primary 4. Finally, it draws on administrative data concerning the child's Primary 1 school (further details are provided in Appendix A).

Table 2-1 provides an overview of the sources of data used in the report. Note that the analysis draws on data for children who took part in language assessments at

[5] More specifically, the results are representative for all children in Scotland in Primary 6 in 2014/15 who were born between June 2004 and May 2005 and who lived in Scotland when they were 10 months old.

both time points only (n = 2944) – that is, children who undertook language assessments both around their 5th birthday and when they were in Primary 6. Furthermore, data were weighted using the GUS longitudinal survey weight, meaning that only cases which have taken part in every face-to-face sweep of GUS up to and including sweep 8 were included. In total, 2726 children were included in the analysis.

Table 2-1 Data overview

Child age / stage and year of data collection	Data sources		
	Parent	**Child**	**Administrative data**
Age 10 months up to 4 years – 2005-09	Information about parent literacy, parental mental wellbeing	None	None
Age 5 (most children in Primary1) – 2009/10	Information about socio-economic characteristics, household measures, child health and development	Objective assessment of child's vocabulary (BAS-II)	
Age 6 (most children in Primary 2) – 2010/11			Data about child's Primary 1 school – consent to linkage obtained from parent when child aged 6
Age 8 (most children in Primary 4) – 2012/13	Information about parenting behaviours, parental engagement and parent-child relationship; significant changes in child's life across primary school years	How child feels about school	
Primary 6 (age 10-11) – 2014/15	Information about socio-economic characteristics, household measures, child health and development; significant changes in child's life across primary school years	Objective assessment of child's vocabulary (WIAT-II)	

2.2. Expressive language ability

Put simply, language development refers to children's use of words, sentences, gestures and vocalisations to convey meaning, communicate with others and gain knowledge (Law et al., 2017). The ability to use language underpins many aspects of

children's activities, including their social interactions and intellectual pursuits, and thereby impacts on various elements of their non-physical development. For example, it contributes to their ability to manage emotions, communicate feelings, form and maintain relationships and read and write. Consequently, as highlighted by Save the Children's (2014) 'Read On. Get On' campaign, solid foundations in early language are the foundation on which children's future education and learning are based. Indeed, a Save the Children study using data from the Millennium Cohort Study (MCS) demonstrated that while socioeconomic disadvantage predicted children's academic performance, 'the most important factor in reaching the expected levels in English and Maths at seven was children's language skills at age five', which was greater than the link to poverty or poor parental education (Finnegan et al., 2015).

The importance of early language development for children's educational attainment has meant that there is also a substantial body of evidence linking low levels of early literacy to poor outcomes in adulthood. For example, research has documented correlations between poor early language development and poor labour market outcomes such as low pay and unemployment (Howieson and Iannelli, 2008; McIntosh and Vignoles, 2001). In addition, a study using data from a UK birth cohort of over 17,000 children born in 1970 found that those with poor vocabulary skills at age 5 were four times more likely to have reading difficulties, three times more likely to have mental health problems and twice as likely to be unemployed by the time they were 34, when controlling for other factors (Law et al., 2009). There is also evidence to suggest that poor early literacy can also be a risk factor associated with criminal behaviour in adulthood (Devitt, 2011; Institute of Education, 2002).

In this report, the focus is on *expressive* language ability, or vocabulary. As outlined above, children's early vocabulary ability has been found to be associated with later outcomes across several parameters (e.g. Law et al., 2009).

In GUS, expressive language ability has been measured three times for children in BC1: when they were just under 3 years old, when they were just under 5 years old and again when they were in Primary 6 (aged 10-11). In this report we focus on the measures obtained at the latter two age points. Across these two age points, the children's vocabulary was measured using two different assessments. These are described below.

As part of the fifth sweep of interviews undertaken with families in BC1 (when the children were aged just under 5 and most were in or about to enter Primary 1), the child's language ability was measured using the naming vocabulary subtest of the *British Ability Scales Second Edition* (BAS-II). This is a cognitive assessment which forms part of the Early Years battery designed for children aged between 2 years and 6 months and 7 years and 11 months. Though numerous tests of language ability

exist, the BAS is particularly suitable for administration in a social survey like GUS. The naming vocabulary test requires the child to name a series of pictures of everyday items to assess their expressive language ability. There are 36 items in total in the assessment. However, to reduce burden and to avoid children being upset by the experience of repeatedly failing items within the scale, the number of items administered to each child is dependent on their performance. For example, one of the criteria for terminating the naming vocabulary assessment is if five successive items are answered incorrectly. As already noted, children in BC1 were asked to complete this assessment when they were just under five years old. As such, the BAS assessment scores offer a snapshot of children's expressive vocabulary ability around the time they started primary school.

When the GUS children were in Primary 6, as part of their eighth GUS interview, their language ability was measured using the listening comprehension subtest of the Wechsler Individual Achievement Test - Second UK Edition (WIAT-II^{UK}). More specifically, the expressive vocabulary measure of the subtest was used, which assesses speaking vocabulary and word retrieval ability. This subtest is part of a comprehensive individually administered test for assessing the achievement of children and adolescents aged between 4 years and 16 years and 11 months. As with the BAS, WIAT is also suitable for administration in a study like GUS, with the version used especially adapted for social surveys. During the expressive vocabulary element of the listening comprehension subtest children are shown a single picture and given an oral description. They then have to provide one word that matches the picture and the description. The assessment has a total of 15 items. As with the BAS, to reduce the burden on the child the number of items administered is dependent on their performance, with the assessment discontinued if the child gives a wrong answer on 6 consecutive occasions. Because children in BC1 were asked to complete this assessment when they were in Primary 6, the WIAT assessment score provides a picture of children's expressive language ability at the time they were nearing the end of primary school.

To make the scores from the two measures of expressive language ability comparable, the vocabulary score from each age point was standardised into a z-score. Z-scores are derived from the survey data. They count the number of standard deviations from the score mean and have a mean of 0. Therefore, a child with a z-score of 0 at either Primary 1 (age 5) or Primary 6 (age 10-11) has an average ability across all children in that age group. Those with a z-score greater than 0 scored above average and those with a score of less than 0 scored below average. The size of the z-score indicates how far above or below average the child's score was. By using the standardised scores, it is possible to compare ability at the two age points and to consider whether children who scored above, below or about

average around the time they entered primary school (aged just under 5) continued to do so when they were in Primary 6.

Throughout the remainder of this report, the terms ‘expressive language ability’, ‘language ability’, ‘expressive vocabulary ability’ and ‘vocabulary ability’ will be used interchangeably.

2.3. Analytical approach and interpreting the results

Much of this report is concerned with exploring expressive language ability for different groups of children according to a number of socio-economic characteristics (annual equivalised household income; highest level of parental education in the household; and area deprivation (SIMD)). Definitions of these measures are provided in Appendix A.

Not all families who initially took part in GUS did so for all subsequent sweeps. There are a number of reasons why respondents drop out from longitudinal surveys and such attrition is not random. Therefore, the data were weighted using specifically designed weights which adjust for non-response and sample selection. All results have been calculated using weighted data and all comparisons take into account the complex clustered and stratified sample structures. Note that because results were calculated using weighted data, the results and bases presented cannot be used to calculate how many respondents gave a certain answer.

Unless otherwise indicated, only differences which were statistically significant at the 95% level or above are commented on in the text.

Notes on how to interpret tables and charts are provided in the text. A brief description of the analysis undertaken is also provided in the text. However, readers interested in the analytical approach should refer to Appendix C.

3 EXAMINING THE GAP IN LANGUAGE ABILITY

3.1. Introduction

This chapter examines the gap in expressive language ability between children from different socio-economic backgrounds at the time they were in Primary 6. The gap is considered according to measures of household income, area deprivation and parental education.

3.2. Expressive language ability by social background - Primary 6

Figures 3-1 to 3-3 display the standardised vocabulary scores of children in Primary 6 by household income, area deprivation and parental education. The distribution of vocabulary scores is shown using box plots (Figures 3-1, 3-2 and 3-3). These are explained below. The numbers informing the box plots are provided in Tables B1 to B3 in Appendix B.

As noted in section 2.2, the scores shown for each socio-economic group are *relative* to those recorded for all children. A score of 0 represents the average vocabulary score recorded for all children, irrespective of social background; a score of -1 represents a score which is one standard deviation below the average for all children and a score of 1 represents a score which is one standard deviation above the average for all children. Children scoring between -1 and 1 make up roughly 70% of all children, while those with scores above 1 and those with scores below -1 make up approximately 15%, respectively.

In each chart, the average (median[6]) score for each socio-economic group is represented by the horizontal line that divides the box into two parts. Half the scores are greater than or equal to this value and half are less. For example, in Figure 3-1, the average vocabulary score for children in the lowest income quintile was -0.25. The box for each socio-economic group represents the middle 50% of scores for that group. The boxes and median lines allow us to compare average vocabulary ability across different socio-economic groups.

The lines extending above and below the boxes – the upper and lower 'whiskers' – represent the range of scores outside the middle 50%. That is, the highest point of the top whisker for each group represents the highest score for children in that particular group while the lowest point of the bottom whisker represents the lowest score for children in that particular group (no outliers were removed). This allows us

[6] The median score is the midpoint of the vocabulary scores recorded for the GUS children – i.e. half of the children will have recorded higher scores than this value, and half will have recorded lower scores than this value.

to consider variations in the full range of ability within each socio-economic group and not just differences in average scores.

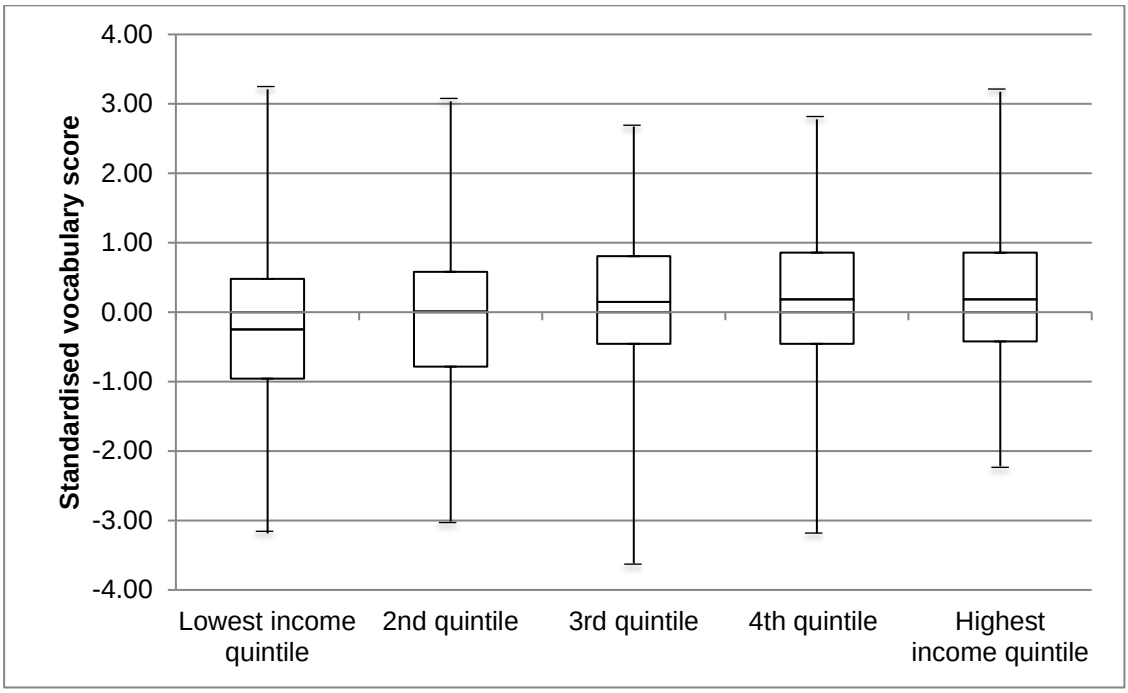
Regardless of the characteristic considered, the graphs show a clear difference in vocabulary ability by social background. In 2014/15, among children in Primary 6, those in higher income households, in less deprived areas, and those whose parents had higher levels of educational qualifications had, on average, better vocabulary than those in lower income households, those in more deprived areas and those whose parent(s) had lower levels of educational qualifications.

The largest differences are visible by parental education. As shown in Figure 3-3, at the time they were in Primary 6, children whose parent(s) had lower Standard Grade qualifications or below had an average vocabulary score of -0.49, compared with 0.25 for children whose parent(s) had a degree. The smallest differences are seen in relation to area deprivation (Figure 3-2), where the average vocabulary score among children in the most deprived areas was -0.29 compared with 0.11 for children in the least deprived areas. Clear differences were also visible by household income (Figure 3-1). Here, the average vocabulary score among children in households in the lowest income quintile was -0.25, compared with 0.18 for children in households in the highest income quintile.

Thus, the charts illustrate a clear gap in language ability by social background when considering average scores for each socio-economic group. Nonetheless, they also illustrate substantial variation *within* socio-economic groups. As noted above, this variation is illustrated by the boxes which mark the middle 50% of scores for each group but also, in particular, by the whiskers at the top and bottom of each box which illustrate the range of ability within each group.

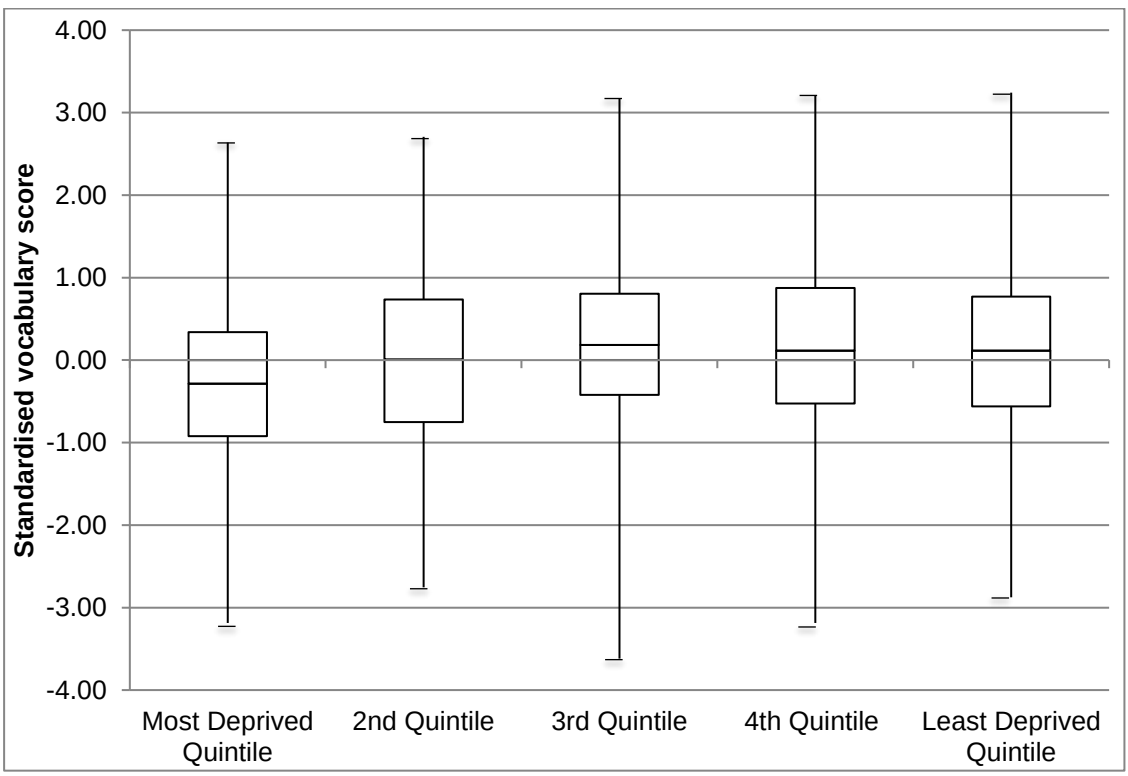
For each of the three social characteristics considered, the charts clearly show that not all children in disadvantaged circumstances did poorly – with significant proportions returning scores above average - and not all children in advantaged circumstances did well – with many having scores below average. It is clear that many children in the least advantaged groups had vocabulary ability as high as or higher than their more advantaged peers whilst some children in the most advantaged groups had poorer ability than some of their less advantaged peers. For example, as shown in Figure 3-1, some children in the bottom income quintile had scores as high as 3.21, higher than any score recorded for children in the highest income quintile.

Figure 3-1 Standardised vocabulary ability score by household income – Primary 6



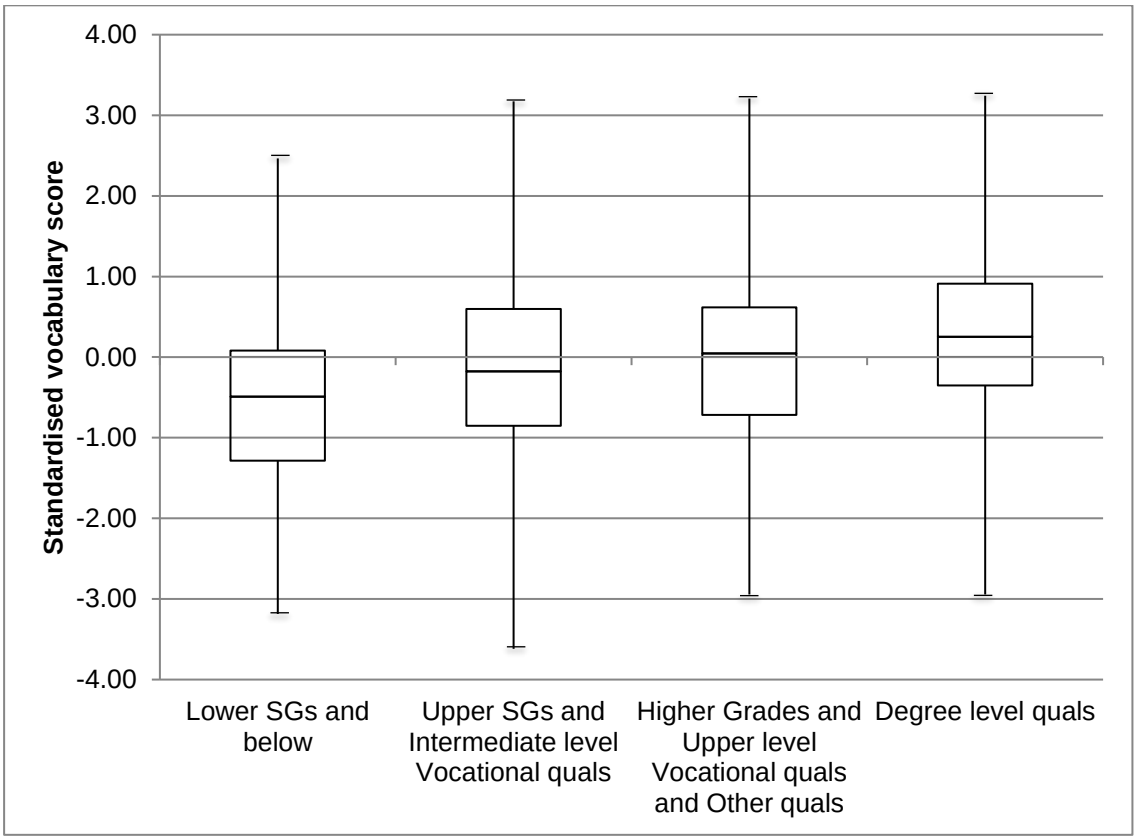
Base size (unweighted/weighted) = 2553/2519. See also Table B-1 in Appendix B.

Figure 3-2 Standardised vocabulary ability score by SIMD – Primary 6



Base size (unweighted/weighted) = 2726/2698. See also Table B-2 in Appendix B.

Figure 3-3 Standardised vocabulary ability score by highest parental level of education – Primary 6



Base size (unweighted/weighted) = 2722/2694. See also Table B-3 in Appendix B.

3.3. Summary

In summary, the analysis showed a gap in expressive language ability between children in the most and least advantaged circumstances towards the end of primary school (in Primary 6). These findings are in line with existing research which has demonstrated a strong correlation between a pupil's socio-economic status and their educational attainment in Scotland, with pupils from disadvantaged backgrounds having a higher risk of not succeeding in school (Scottish Government, 2017a).

Similar to an earlier GUS report which examined the gap in cognitive ability over the pre-school period (Bradshaw, 2011), the largest gap was seen in relation to parental level of education, with smaller but still substantial gaps according to household income and level of area deprivation.

Notably, though, although there were clear differences in average language performance between children in different socio-economic groups, there was also significant variation within these groups. Despite the poorer average performance of children from the most disadvantaged background, some of these children were performing well. Conversely, despite a better average performance, some children in the most advantaged groups were performing less well than some children in the

most disadvantaged groups. These variations indicate that social background, whilst an important factor, is not the only factor that influences language ability. Although being from a disadvantaged social background increases the risk of poorer language skills, it does not equate to poorer language skills for all children from disadvantaged backgrounds.

4 FACTORS ASSOCIATED WITH IMPROVEMENT

4.1. Introduction

This chapter looks at factors present in children's lives over the primary school years which may help or hinder their language development. It seeks to identify characteristics and circumstances that might help to improve children's language ability relative to their peers.

A previous GUS report (Bradshaw, 2011) demonstrated that children's vocabulary ability differs according to social background upon entry to primary school. In a similar vein, the previous chapter showed clear differences in vocabulary ability according to a number of social background characteristics towards the end of primary school (in Primary 6). However, the analysis undertaken thus far does not allow us to explore whether social background is associated with a relative change in ability over the period – in other words, whether children from less advantaged backgrounds improve at a faster, similar or slower rate than children from more advantaged backgrounds. It also does not allow us to determine whether each characteristic is associated with language ability independently of the other characteristics. For example, it is unclear whether the differences seen by income are driven by other differences among children from different income groups such as the parents' level of education, their parenting practices, and/or the child's experience at school.

The emphasis in this chapter is on factors associated with this *relative change* in language ability over the primary school period, rather than with ability at a single time point. After outlining the key factors considered in the analysis, the chapter briefly considers the relationship between expressive vocabulary ability at the two time points considered in the analysis, namely around the time the GUS children started school and again when they were in Primary 6. Then, drawing on univariate linear regression models fitted for several characteristics, circumstances and experiences (outlined in Table 4-1 below), it explores which (if any) are associated with a relative change in language ability between the two time points – that is, over the primary school years. Next, it draws on multivariable regression models to explore which of the factors found to be associated with a relative change in expressive language in the initial analysis *remain* associated with a relative improvement or decline in ability once other known differences between the children are taken into account. Finally, the chapter explores whether any associations found vary according to the children's social background – more specifically, according to

their parents' level of education. This is done through the fitting of interaction effects to the final multivariable regression model[7].

4.2. Factors considered in the analysis

The analysis considers a range of different factors which are known or expected to be associated with children's language development. These are listed in Table 4-1 below; further details are provided in Appendix A[8].

The factors explored in the analysis have been selected for the following combination of reasons: existing research has shown associations with children's language ability; they are considered likely to play an important role in children's lives over the primary school period and thus be (directly or indirectly) important for their language development and GUS has collected data suitable for exploring them.

[7] Further details of the analysis undertaken are provided Appendix C.

[8] Note that where variables have been banded, this has been done to create varied categories large enough to support the analysis whilst reflecting the variation of the responses within the full variable.

Table 4-1 Characteristics, circumstances and experiences considered in the analysis

Variable	Age/stage of child when measured	Source (parent/child/administrative data)
Child's gender	Age 10 mths	Parent
Social background and location		
Highest level of parent education	Age 5/P1	Parent
Annual household income	Age 5/P1	Parent
Level of area deprivation (SIMD)	Age 5/P1	Parent
Urban vs small town or rural location	Age 5/P1	Parent
Other household factors		
Whether languages other than English spoken in the household	Age 5/P1	Parent
Whether parent reported any literacy issues	Age 4	Parent
Parent mental wellbeing	Age 5/P1 and P6	Parent
Whether parent has limiting long-term health problem	Age 5/P1 and P6	Parent
Child health and development		
Whether child has a limiting long-term health problem	Age 5/P1 and P6	Parent
Whether child has above average levels of social, emotional and behavioural difficulties	Age 5/P1 and P6	Parent
Significant life event occurred		
Whether child experienced parental separation or re-partnering	Age5/P1, Age 6/P2, Age 8/P4 and P6	Parent
Whether child changed school	Age5/P1, Age 6/P2, Age 8/P4 and P6	Parent
Whether child experienced adverse life event (death of a parent or sibling; a parent in prison; being in care; or a parent losing their job)	Age5/P1, Age 6/P2, Age 8/P4 and P6	Parent
School		
Child's feelings about school	Age 8/P4	Child
Size of school	Age 5/P1	Administrative data
Whether denominational school	Age 5/P1	Administrative data
Proportion of children registered for free school meals	Age 5/P1	Administrative data
Parenting and parent-child relationship		
Warmth of parent-child relationship	Age 8/P4	Parent
Parental interactions with child's school	Age 8/P4	Parent
How often parent helps child look for school-related information	Age 8/P4	Parent
Child home reading in last week	Age 8/P4	Parent
Parent belief that they can influence child's achievements at school	Age 8/P4	Parent

Including a measure of the child's gender in the analysis allows us to examine whether there are differences in the level at which boys' and girls' vocabulary

improves (or declines), relative to their peers', during the first five years after they start school.

As in chapter 3, household income, area deprivation and parental education are considered as measures of social background. To ensure sufficient base sizes for the type of analysis undertaken, rather than break parents' level of education into four groups as was done in chapter 3, the measure of parental education used in this chapter simply identifies whether or not at least one parent or carer in the household was educated to degree level (or above).

A measure of whether the child resided in an urban, small town or rural area provides a means to explore whether geographical location appears to be associated with children's language development. Any such association may arise through differences in general lifestyle, but may also, for example, arise through differences in school experiences in urban, small town or rural communities, as the size, resources and ethos of schools based in these different areas are likely to differ (e.g. Commission on the Delivery of Rural Education, 2013). To support the analysis, areas in the Scottish Government's six-fold urban-rural classification[9] were grouped to form a binary variable comparing children living in large and other urban areas with those living in small town or rural areas (see Appendix A).

Bi- and poly-lingual children can have slightly delayed language development, but then catch up with their peers as they grow older (Cattani et al., 2014). Exploring whether languages other than English are spoken in the household at the time the child enters primary school allows us to gauge whether children in bi- or poly-lingual homes appear to be more or less likely to see an improvement in their language skills during the first years of primary school, relative to their mono-lingual peers. The analysis also considers parental literacy, as a parent experiencing literacy issues may impact negatively on their child's language development. For example, parents who have literacy issues may be less inclined to read with their children or engage in other educational activities which may help improve their child's language skills.

Other household factors such as the parent's mental wellbeing, and whether they had a limiting long-term health problem are also explored. Parental mental health and wellbeing has been shown to be strongly associated with children's cognitive development at an early stage (e.g. Marryat and Martin, 2010; Barnes et al, 2010). Compared with parents who have higher levels of mental wellbeing, someone who suffers from poor mental health may have less energy and/or capacity to engage in activities with the child known to improve language development. Parents or carers who have a long-term limiting health problem may be similarly inhibited in their parenting activities and/or there may be further impacts on the home environment

[9] https://www2.gov.scot/Topics/Statistics/About/Methodology/UrbanRuralClassification

such as financial constraints resulting from lower earnings and/or additional costs associated with their health problem.

Children's health and social development are also considered following other evidence linking these to cognitive outcomes (e.g. Gregg and Washbrook, 2011). Having a limiting health problem or social development difficulty may severely affect a child's language development in several ways, depending on the issue in question.

Experiencing significant changes or events can have a substantial effect on children and may, directly or indirectly, influence their language development. For example, changing school often carries with it not only changes in surroundings and staff but also in peer relationships, and may cause a general sense of upheaval which is potentially detrimental to children's learning. Changes in their home environment caused by parents separating and/or step-parents moving in has also been shown to be associated not just with children's wellbeing, but also with their cognitive outcomes (Chanfreau et al., 2011). Furthermore, experiencing an event such as the death of a parent or sibling; a parent being in prison; spending time in care; or a parent losing their job is likely to cause upset and distress which could have indirect effects on children's learning, for example through being off school for a prolonged period or simply being unable to concentrate. The analysis considers three measures related to significant changes in the child's life over the primary school period: whether they experienced parental separation or re-partnering; whether they changed school; and whether they experienced a significant adverse life event such as the death of a parent or sibling, a parent being in prison, spending time in care or a parent losing their job.

In this report we are particularly interested in factors which are potentially modifiable in the short to medium term through relatively discrete initiatives. That is, factors which can be influenced through dedicated changes to policy or practice either at a national level or through targeted interventions and initiatives. As such, aspects of the child's schooling and parenting practices are of particular interest here.

As noted in section 1.3, current education policies in Scotland highlight the importance of children's mental wellbeing for their academic achievement. How a child feels about school is likely to impact on their learning experience and the way in which they engage in school activities – and, ultimately, on their learning outcomes. Conversely, a child's skills and abilities are also likely to influence how they feel about school. Either way, it is useful to understand how, if at all, children's feelings about school may be associated with aspects of their cognitive development – including their expressive language ability – as they progress through school.

Understanding more about which (if any) aspects of a child's school experience are associated with higher levels of improvement is something which is of obvious interest to policy makers and may help focus both policy making processes and further research efforts. The analysis considers several school-related measures taken from both survey and administrative data:

- A measure of the cohort **child's enjoyment of school** around the time they were in Primary 4 (aged just under 8 years), based on the child's own report.
- **Information about the child's Primary 1 school**, obtained through linkage with administrative records. Specifically:
    - the size of the school;
    - whether the school was denominational or not;
    - the proportion of children in the school who were registered for free school meals (25% or more compared with less than 25%)[10].

Children attending large primary schools are likely to have very different experiences than those in the smallest schools, with potential benefits and drawbacks of each. The experience of attending a denominational school may also differ to that of attending a non-denominational school. Furthermore, in statistics of school performance, denominational schools have been found to perform particularly well (see e.g. Andrews and Johnes, 2016; Hinchliffe and Bradshaw, 2015). As for the proportion registered for free school meals, this is a commonly used indicator of the level of poverty among the pupils attending the school. In Scotland, eligibility for free school meals has now been extended to all children in Primary 1 to Primary 3, however, at the time the children in BC1 were in Primary 1, eligibility for free school meals was still determined based on need, primarily though the family's eligibility for and receipt of certain benefits. These measures are obviously not a comprehensive set of indicators of the child's school experience and for a significant minority of cases no information was available on measures of school size, denomination or the proportion registered for free school meals. Nonetheless, they do provide measures of selected aspects which may, if nothing else, suggest possible fruitful directions for future research.

The relationship between parenting and children's development has received much attention in recent years and existing research has shown numerous links between a range of parenting and home learning activities and children's cognitive development (e.g. Waldfogel and Washbrook, 2010; Bromley, 2009; Melhuish, 2010). Among policy makers there has been a particular interest in measuring and encouraging 'parental engagement' in their child's school and education. These efforts are targeted at parents, encouraging them to engage in educational activities with their children at home, as well as at schools and teachers to ensure they are maximising the opportunities for parents to be meaningfully involved in their child's schooling (e.g. Scottish Government 2017b).

The analysis considers the following measures of parenting and home learning activities and characteristics – all are based on data collected from the cohort child's

[10] GUS data were collected in 2009/10 before free school meals were rolled out to all children in Primary 1 to Primary 3 introduced in 2015.

parent around the time the child was aged just under 8 years old, i.e. when most children were in Primary 4[11]:

- **Warmth of the parent-child relationship**
- **Parental engagement** in child's schooling, including:
    - parent interactions with the child's school (those with 7 to 10 different types of different interaction compared with those with 6 or less);
    - how often a parent helped the child look for information about what he/she was learning at school, for example at the library or on the internet;
    - a measure of the extent to which the parent believes they can influence their child's achievements at school (those holding very positive beliefs compared with those holding less positive, neutral or negative beliefs).
- **Home learning activities**: how many days in the last week the child read or looked at books at home (children doing so at least 6 days per week compared with those doing so less often).

The measure of parent interactions with the school is a count of the number of different activities the parent reported to have engaged in in the two years before the interview[12].

4.3. Vocabulary ability upon entry to primary school and in Primary 6

Before considering which factors may help or hinder a relative improvement in language ability, it is worth exploring the relationship between the two measures of vocabulary ability used in the analysis – that is, the measure obtained around the time the children started primary school, and then when they were in Primary 6.

As shown in Table 4-2, we see a strong relationship between the two standardised scores, with around 17% of the variation in standardised vocabulary scores in Primary 6 explained by the variation in scores at the start of primary school. In other words, a substantial proportion of the differences in children's expressive language ability at the time they are in Primary 6 appears to be explained by their ability around the time they started school. This also indicates that children's language ability at primary school entry is closely related to their ability towards the end of primary school. Nonetheless, the proportion of variation explained is not as large as has been found in some other studies (see e.g. Goodman, Gregg and Washbrook, 2011), and a rather large proportion of the variation in language skills at the time children were in

[11] Full details are provided in Appendix A.

[12] Around the time the cohort children turned 8 years, when most children were in Primary 4, parents were asked about the period since their last GUS interview which was undertaken when the cohort child was aged just under 6 years, i.e. when most children were in Primary 2. Further details are provided in Appendix A.

Primary 6 does *not* appear to be explained by their earlier ability – at least the way it is measured here. On this point, it is worth bearing in mind that the analysis uses two different measures of expressive vocabulary, something which (despite the use of 'standardised' scores, as outlined in section 2.2) is likely to have introduced higher levels of uncertainty in the analysis than if the exact same measures had been used at both time points. Even with this caveat, however, the results suggest that although children's expressive language ability around the time they start primary school appears to play an important role in explaining their level of ability towards the end of primary school, other factors are also important.

Table 4-2 Standardised expressive vocabulary score in Primary 6, by standardised expressive vocabulary score in Primary 1

	Standardised expressive vocabulary score (Primary 6)			
			95% Confidence interval***	
	p-value*	Regression coefficient**	Lower	Upper
Standardised expressive vocabulary score (Primary 1)	.000	0.431	0.385	0.476
R squared	0.172			
Weighted base	*2698*			
Unweighted base	*2726*			

* All figures quoted in this report have a margin of error because they are estimates based on a sample of children, rather than all children. The p-value is an estimation of how likely it is that we would find a relationship in our sample of children if there was no actual relationship in the population (i.e., broadly speaking, among children in Scotland who are the same age as the GUS children but who are not part of GUS). Thus, the smaller the p-value ($p<0.05$), the more confident we can be that our results are likely to apply to children in Scotland more widely.

** The regression 'coefficient' illustrates the relative level of change (positive or negative) in language ability score at P6 if score at P1 is increased by 1 unit. A significant ($p<0.05$) positive coefficient denotes a relative improvement in ability score and a significant negative coefficient denotes a relative decline in ability score for every one-unit increase in P1 score.

*** The 95% confidence interval is an indication of the level of uncertainty in the coefficient estimate.

4.4. Univariate regression analysis

The following sections explore the extent to which a range of characteristics, circumstances and experiences are associated with a relative change in vocabulary ability over the primary school period.

The first stage of the analysis considers the relationship between each factor – that is, each of the characteristics, circumstances and experiences set out in Table 4-1 – and a relative change in language ability between the start of primary school and Primary 6. This is done by fitting separate linear regression models for each factor with the standardised Primary 6 vocabulary score as the dependent (outcome) variable and the Primary 1 standardised vocabulary score as an additional independent variable (co-variate). This allows us to assess the relationship between

each factor of interest and a relative change in language ability between the two time points[13].

Table 4-3 lists the factors which the analysis showed to be associated with a relative change in language ability over the period (when other differences between the children are not controlled for). Only factors where the relationship with a relative change in vocabulary ability is statistically significant at the 10% level are included in the table[14]. A '+' in the 'Direction of change' column indicates a positive relationship between having the characteristic in question and a relative improvement in expressive language ability; conversely, a '-' indicates a negative relationship and a relative decline in language ability. For example, the '+' for 'Higher household income' indicates that children in higher income households improved at a higher rate than children in low income households (the reference category).

Table 4-3 Factors individually associated with a relative change in expressive vocabulary ability (not controlling for other differences)

Factor	Direction of change
Social background and geographical location	
Higher income household (vs low income)	+
In less deprived area (vs most deprived)	+
Parent/carer educated to degree level (vs not degree educated)	+
Live in small town or rural area (vs urban)	+
Child development	
Child has above average social, emotional or behavioural difficulties (vs average or below)	-
School	
Child's primary 1 school has high % of children eligible for free school meals (25% or more vs less than 25%)	-
Parenting	
Parent reads with child at least 6 days per week (vs less often)	+
High number of parent interactions with school (7-10 vs 6 or less)	+
Parent holds strong positive belief they can influence child's achievement at school (vs less positive, neutral or negative beliefs)	+

[13] Only children with valid vocabulary scores at both time points were included in the analysis (36 children with a valid vocabulary score at age 5/Primary 1 were excluded from the analysis because there was no valid vocabulary score at Primary 6). Furthermore, data were weighted using the GUS longitudinal survey weight, meaning that only cases which have taken part in every face-to-face sweep of GUS up to and including sweep 8 were included. In total, 2726 children were included in the analysis.

[14] Reducing the statistical boundary to 10% allowed the inclusion of a small number of factors where the significance value was close to the typical 5% level. Full analytic outputs for the linear regression models for each of the factors listed in Table 4-1 are provided in Appendix C.

Table 4-3 shows that over the course of the primary school period, children in more advantaged circumstances improved at a higher rate than their less advantaged peers. In relation to income, the analysis shows a clear positive relationship with children in the higher income groups improving at a higher rate than those in the lowest income group. Meanwhile, children living in less deprived areas were more likely to improve than those living in the most deprived areas. Those with degree-educated parent(s) improved at a higher rate than those whose parent(s) did not have a degree and children living in small town or rural areas showed higher levels of improvement compared with their peers living in urban areas.

Looking at the child's social development, those who were reported by their parent as having above average levels of social, emotional and behavioural difficulties saw lower levels of improvement in their vocabulary ability, compared with those who had only average or below average levels of difficulties. None of the other measures of household factors nor the child's health or significant changes in the child's life were found to be associated with a relative change in language ability over the period. Neither were the child's feelings about school, the school size, nor school denomination. However, children who attended a school with more than 25% of pupils registered for free school meals were more likely to see a relative decline in their expressive language skills compared with those attending a school with a lower proportion registered for free school meals.

Three measures of parenting were individually associated with a relative improvement in vocabulary. First, children whose parents read with them at least 6 days a week around the time they were in Primary 4 saw higher levels of improvement compared with children whose parents read with them less often. Second, children whose parents reported a high number of interactions with the child's school improved at a higher rate than children whose parents reported fewer interactions with the school. Third, children whose parents held strong positive beliefs in their ability to influence the child's achievement at school improved more than children whose parents held less positive beliefs.

As already noted, for each of the factors explored, the analysis carried out here did not take into account other differences between the children. Thus, the relationships outlined above may have arisen because of other differences between the children particularly given how these factors tend to vary by social background. This question is explored as part of the multivariable analysis outlined below.

4.5. Multivariable regression analysis

The next stage of the analysis involved entering the statistically significant factors listed in Table 4-3 above, as well as the child's gender, into a single multivariable regression model. As for the regression models outlined above, standardised Primary 6 vocabulary score is included as the dependent (outcome) variable and Primary 1 standardised vocabulary score as an independent variable (co-variate). This

approach allows us to explore the extent to which each factor remains associated with a relative change in ability once these other known differences are controlled for.

This analysis showed[15] that only four factors remained associated with a relative change in vocabulary ability once other factors had been controlled for:

- **Parental education** – children with parent(s) educated to degree level or above saw higher levels of improvement than those whose parent(s) had lower levels of education.
- **Urban or small town/rural location** – children living in small town or rural areas improved more than those in urban areas.
- **Child's level of social, emotional and behavioural difficulties** – children with above average levels of difficulties were more likely to see a relative decline in ability than their peers with average or below average levels of difficulties.
- **Home reading** – children who read or looked at books at home at least six days per week improved at a higher rate than those who looked at books less frequently.

A final model was fitted with only those factors which were statistically significant in the multivariable analysis described above, as well as the child's gender. The results are outlined in Table 4-4 below.

4.6. Variation by parental education

The final stage in the analysis explored whether any of the associations found vary according to the parents' level of education. This was done through fitting so-called 'interaction effects' between parental education and each of the remaining variables to the final regression model outlined in Table 4-4[16]. None of the interaction effects were statistically significant. This suggests that the relationship between a relative change in ability and each of the factors identified at the earlier stages of the analysis and outlined above – i.e. urban or small town/rural location; the child's level of social, emotional and behavioural difficulties; and home reading – does not vary systematically according to parents' level of education.

[15] Full results are provided in Appendix C.

[16] Details of the analysis undertaken are provided in Appendix C.

Table 4-4 Factors associated with a relative change in vocabulary ability (controlling for other differences) – final model

	Change in expressive vocabulary ability			
	P-value	Coeff*	95% confidence interval	
			Upper	Lower
Child's gender (ref=Girl)				
Boy	.119	.063	-.017	.142
Highest parental level of education (ref=Below degree)				
Degree	.001	.143	.061	.225
Location (ref=Urban)				
Small town or rural	.013	.124	.027	.220
Child level of social, emotional, behavioural difficulties (ref=Average or below)				
Above average level of difficulties at one or both time points	.002	-.179	-.290	-.068
Home reading in last week (ref=5 days or less)				
6-7 days	.012	.118	.027	.209
Unweighted base	*2726*			
Weighted base	*2698*			

* In this table the regression 'coefficient' illustrates the relative level of difference (positive or negative) in language ability for each sub-group as compared with the reference sub-group. A significant ($p<0.05$) positive coefficient denotes a greater improvement in ability score and a significant negative coefficient denotes a lower change in ability score when compared with the reference sub-group. The reference sub-group is indicated in brackets.

4.7. Summary

This chapter has shown that, on average, when considered on an individual basis – that is, when not taking into account any other known differences – children living in higher income households, those living in areas with lower levels of deprivation, and those with parent(s) educated to at least degree level saw higher levels of improvement in their expressive language ability relative to their peers than those in the lowest income households, those in the most deprived areas, and those whose parent(s) do not have a degree.

Furthermore, it identified four experiences and circumstances which were associated with a relative improvement in children's expressive vocabulary over the primary school years even after a range of other known differences were controlled for:

- Having parent(s) educated to degree level or above
- Living in a small town or rural area

- *Not* having above average levels of social, emotional or behavioural difficulties
- Reading or looking at books at home at least 6 days a week when aged 8/in Primary 4

The analysis showed no indications that these relationships vary systematically according to parental education, suggesting that these factors are associated with improvement for all children irrespective of whether their parent or parents are educated to degree level or not.

5 SUMMARY AND CONCLUDING REMARKS

5.1. Introduction

This chapter draws together the analysis presented in the previous chapters to answer the research questions set out in section 1.5. It also suggests some implications for policy makers and others seeking to improve attainment for children in Scotland.

5.2. Does the gap in expressive language ability between children from advantaged and disadvantaged backgrounds change over the primary school years?

Other research has found that educational inequalities are exacerbated as children move up through the school system (e.g. Sosu and Ellis, 2014; Scottish Government, 2017b). In a Scottish context, we know from previous GUS research (Bradshaw, 2011) that inequalities in expressive language ability exist upon entry to primary school, with less advantaged children already falling behind their more advantaged peers.

If this gap was narrowing, in the analysis carried out here we may have expected to see children in less advantaged circumstances improving at a higher rate than their more advantaged peers. However, as outlined in chapter 4, findings from the analysis do not provide any evidence of this happening. On the contrary, they suggest that children living in higher income households, children in less deprived areas, and children with parent(s) educated to degree level improved *more*, relative to their peers, than those in the lowest income households, those in the most deprived areas, and those whose parents did not have a degree, respectively.

It is important to note that the analysis carried out here focusses predominantly on identifying factors associated with helping or hindering improvement (see below), rather than on measuring the size of the attainment gap. Nonetheless, the results do seem to suggest, if anything, that inequalities in expressive language ability appear to have widened rather than narrowed over the primary school period.

As demonstrated in chapter 3, the report has shown that the gap between more and less advantaged children seen in previous GUS research persists and is evident as children reach the last years of primary school – irrespective of whether we measure the gap in relation to differences by family income, the level of area deprivation or parental level of education. In line with earlier research (e.g. Bradshaw, 2011), the largest differences in ability were seen in relation to parental education, with smaller

but still significant gaps evident according to family income and the level of area deprivation.

Having said this, alongside demonstrating clear inequalities between groups of children, the analysis also showed substantial levels of variation in ability *within* the social groups considered. These variations indicate that social background, whilst an important factor, is not the only driver influencing language ability. Although being from a disadvantaged social background increases the risk of poorer language skills, it does not equate to poorer language skills for *all* children from disadvantaged backgrounds. Similarly, coming from a more advantaged background does not guarantee more advanced language development.

5.3. What circumstances and experiences are associated with a relative change in expressive language ability over the primary school years?

As already noted, raising educational attainment is an important priority for the Scottish Government. A key question, then, is what might support children to do (even) better, irrespective of their social background? With an emphasis on expressive language development, this report has identified several characteristics and circumstances which appear to be associated with children showing greater improvement in their language skills relative to their peers.

One of the characteristics found to be associated with a relative change in language ability was the extent to which children were reported as having social, emotional or behavioural difficulties; compared with those with no difficulties, children with higher than average levels of social development difficulties saw a relative decline in ability over the primary school period. Whilst the analysis did not consider different aspects of development in detail, other research has suggested that hyperactivity and conduct problems appear to play a role in relation to educational attainment towards the end of primary school (see e.g. Gregg and Washbrook, 2011). For policy makers and others involved in supporting children's learning, this highlights the need for policies and initiatives aimed at supporting children's educational attainment to take into account other aspects of the child's development too. This finding emphasises the importance of ensuring that children with additional support needs associated with social and behavioural development are fully supported during their primary school education. Encouragingly, this is already to some extent recognised through the emphasis on health and wellbeing in both the National Improvement Framework and CfE.

Looking at broader circumstances, living in a small town or rural was predictive of a relative improvement in language ability. This relationship remained statistically significant even when controlling for other known differences between the children, including parental education and aspects of the home learning environment. It is not clear from the analysis carried out here what may explain this association between living in a small town or rural area and a relative improvement in language ability. However, previous GUS reports, among others, have demonstrated the importance

of the quality of children's pre-school setting for their cognitive development up until the start of primary school (Bradshaw et al., 2014; Knudsen et al., 2017; see also Scobie and Scott, 2017 for an overview of the literature). It is not unlikely that the quality of the school setting also has a bearing on children's language development – and that the level of quality varies according to the location of the school. Equally, however, the experiences of children in small town or rural areas may differ in numerous other ways, including on a range of lifestyle measures. Further research would be useful to understand more about the differences in urban and small town/rural experiences which may be important for language development as well as wider attainment.

The report has also demonstrated a positive relationship between frequent reading at home when the child was aged 8 and a relative improvement in language ability over the primary school period. This may simply indicate that children who experience an improvement in their language ability are (or become) more likely to show an interest in reading and are thus more likely to read at home. Nonetheless, home learning activities, including frequent reading in the early years, has previously been shown to be associated with a relative improvement in educational attainment between the ages of 7 and 11 (Gregg and Washbrook, 2011), and previous GUS research showed associations between early parent-child reading and language ability at ages 3 and 5 (Bromley, 2009; Bradshaw, 2011). Along these lines, the relationship seen here appears to suggest that home learning activities – and specifically reading – continue to play a role for children's language development beyond the early years. This provides some support for the rationale behind initiatives such as the Scottish Government's 'Read, Write, Count' campaign which encourages parents of early primary school-aged children to engage in educational activities with their child to support their learning, and Save the Children's 'Read On, Get On' campaign (both initiatives are described in more detail in section 2.2).

It is worth noting that the factors included in our analysis and which emerge as important only explain a limited amount (around 19%) of why some children show greater improvement in their language ability over the primary school period.

5.4. Do the factors associated with a relative change in ability vary according to social background?

The analysis found no indications that the characteristics and circumstances found to be associated with a relative change in expressive language ability varied according to children's social background (measured here through parental level of education).

5.5. Concluding remarks

This report has demonstrated that the gap in language ability between the most and least socio-economically advantaged children evidenced around the time they started school (Bradshaw, 2011) was still very much evident by the time they reached Primary 6. The findings also appear to suggest that inequalities widened over the

primary school years, although the analysis undertaken does not allow us to estimate by what margin the gap has widened.

Notably, the report has also demonstrated how, despite clear inequalities in *average* vocabulary ability among children in the most and least advantaged groups, there was a large amount of variation in ability within socio-economic groups. Some children from disadvantaged backgrounds were doing well relative to their more advantaged peers whilst, conversely, some children in advantaged circumstances were doing less well than their less advantaged peers. What this seems to suggest is that although being from a disadvantaged social background increases the risk of poorer language skills, it does not equate to poorer language skills for *all* children from disadvantaged backgrounds. Similarly, not all children from more advantaged backgrounds necessarily have better language skills. Thus, support for children should not operate solely on the basis of socio-economic characteristics when considering children and young people's barriers to learning.

Importantly, the report also identified a small number of characteristics and circumstances which were associated with children demonstrating either higher or lower levels of improvement over the primary school period, relative to their peers. These factors only explain a limited amount of why some children show greater improvement in their language ability than others over the primary school period. This means there are a range of additional characteristics and circumstances which must be considered to fully address inequalities in language development. Nevertheless, some factors which were associated with improvement are worth noting.

First, children with above average levels of social, emotional and behavioural difficulties were at a disadvantage – these children were less likely to improve relative to their peers than children with lower levels of difficulties. This finding appears to lend support to efforts and initiatives that take a holistic approach to considering children's attainment which does not narrowly focus on one aspect (such as their language), but sees this as part of their wider development, including their mental health and wellbeing. It also emphasises the importance of ensuring that children with social and behavioural developmental needs are properly supported during their primary education.

The report also showed higher levels of improvement among children in small town and rural areas, even when other known differences such as parental education were taken into account. This seems to suggest that there are systematic differences in children's experiences in and/or outside of education in the areas where they live which affect their language development, and which we were not able to take into account here. These may be, for example, differences in lifestyle, in the quality of the school environment, and/or the quality of teaching. Further research to better understand the drivers of these differences would be useful.

Finally, the report showed higher levels of improvement relative to other children among those who read or looked at books at home every day or almost every day

around the time they were 8 years old (for most children, when in Primary 4), irrespective of other known differences such as the parent's level of education. Whilst this may reflect that children who experience an improvement in their language ability develop (more of) an interest in reading, it may also be an indication that home learning activities – and reading in particular – is beneficial for children's language development beyond the early years, thus lending some support to campaigns encouraging parents to continue to engage in home learning activities with their child after they have started school, as well as campaigns aiming to encourage reading among children more widely.

6 REFERENCES

Andrews, J. and Johnes, R. (2016) F*aith Schools, Performance and Social Selection.* London: Education Policy Institute.

Barnes, M. Chanfreau, J., Tomaszewski, W. (2010) *Growing Up in Scotland: The circumstances of persistently poor children.* Edinburgh: Scottish Government.

Bradshaw, P. (2011) *Growing Up in Scotland: Changes in child cognitive ability in the pre-school years.* Edinburgh: Scottish Government.

Bradshaw, P., Lewis, G., and Hughes, T. (2014) *Growing Up in Scotland: Characteristics of pre-school provision and their association with child outcomes*. Edinburgh: Scottish Government.

Bromley, C. (2009) *Growing Up in Scotland: Year 3 – The Impact of Children's Early Activities on Cognitive Development.* Edinburgh: Scottish Government.

Cattani, A., Abbot-Smith, K., Farag, R., Krott, A., Arreckxy, F., Dennis, I. and Floccia, C. (2014) How much exposure to English is necessary for a bilingual toddler to perform like a monolingual peer in language tests? In International Journal of Language and Communication Disorders, vol. 49, no. 6, pp. 649-671.

Chanfreau, J., Barnes, M., Tomaszewski, W., Philo, D., Hall, J. and Tipping, S. (2011) *Growing Up in Scotland: Change in early childhood and the impact of significant events*. Edinburgh: Scottish Government.

Commission on the Delivery of Rural Education (2013) *Report form the Commission on the Delivery of Rural Education*. Edinburgh: Scottish Government.

Devitt, K. (2011) *Young Adults Today: Education, training and employment and young adults in the criminal justice system fact file.* Brighton: Young People in Focus.

Finnegan, J., Telfer, C. and Warren, H. (2015) *Ready to Read: Closing the gap in early language skills so that every child in Scotland can read well.* London: Save the Children.

Goodman, R. (1997) The Strengths and Difficulties Questionnaire: a research note, *Journal of Child Psychology and Psychiatry*, 38, 581-586.

Goodman, A., Gregg, P. and Washbrook, E. (2011) 'Children's educational attainment and the aspirations, attitudes and behaviours of parents and children through childhood in the UK'. *Longitudinal and Life Course Studies*, 2(1), pp. 1-18.

Gregg, P and Washbrook, E (2011) 'The role of attitudes and behaviors in explaining socio-economic differences in attainment at age 11', *Longitudinal and Life Course Studies*, 2(1), pp. 41-58.

Hinchliffe, S. and Bradshaw, P. (2015) *The identification of effective secondary schools in challenging circumstances: Summary report.* Edinburgh: Scotland's Commissioner for Children and Young People.

Howieson, C. and Iannelli, C. (2008) 'The effects of low attainment on young people's outcomes at age 22-23 in Scotland', *British Educational Research Journal*, 34(2), pp. 269–90.

Institute of Education (2002) *Basic Skills and Crime: Findings from a study of adults born in 1958 and 1970*. London: The Basic Skills Agency.

Knudsen, L; Currie, E & Bradshaw, P (2017) *Changes in Early Learning and Childcare Use and Outcomes at Age 5: Comparing Two Growing Up in Scotland Cohorts*. Edinburgh: Scottish Government.

Law, J., Rush, R., Schoon, I. and Parsons, S. (2009) 'Modelling developmental language difficulties from school entry into adulthood: literacy, mental health, and employment outcomes', *J Speech Lang Hear Res,* 52(6), pp. 1401-16.

Law, J., Charlton, J and Asmussen, K. (2017) *Language as a child wellbeing indicator.* London: Early Intervention Foundation.

Marryat, L. and Martin, C. (2010) *Growing Up in Scotland: Maternal mental health and its impact on child behaviour and development.* Edinburgh: Scottish Government.

Melhuish, E. (2010) *Growing Up in Scotland: Impact of the home learning environment on child cognitive development.* Edinburgh: Scottish Government.

McIntosh, S. and Vignoles, A. (2001) Micro-analysis of the Effects of Literacy and Numeracy. In: Bynner et al. *Improving Adult Basic Skills: Benefits to the individual and to society.* London: Centre for Longitudinal Studies.

OECD (2015) *Improving Schools in Scotland: An OECD Perspective.* Paris: OECD.

Ritchie, S. J., Bates, T. C. and Polmin, R. (2015) Does learning to read improve intelligence? A longitudinal multivariate analysis in identical twins from age 7 to 16, *Child Development,* 86(1), pp. 23-36.

Scobie, G. and Scott, E. (2017) *Rapid evidence review: Childcare quality and children's outcomes*. Edinburgh: NHS Health Scotland.

Scottish Government (2008) *Curriculum for Excellence: Building the Curriculum 3.* Edinburgh: Scottish Government.

Scottish Government (2016) *Summary statistics for attainment, leaver destinations and healthy living No. 6: 2016 Edition.* Edinburgh: Scottish Government.

Scottish Government (2017a) *Scottish Survey of Literacy and Numeracy 2016 (Literacy).* Edinburgh: Scottish Government.

Scottish Government (2017b) *2018 National Improvement Framework and Improvement Plan for Scottish Education.* Edinburgh: Scottish Government.

Scottish Government (2018) Getting It Right for Every Child, http://www.gov.scot/Topics/People/Young-People/gettingitright.

Sosu, E. and Ellis, S. (2014) *Closing the Attainment Gap in Scottish Education*. York: Joseph Rowntree Foundation.

Waldfogel, J., and Washbrook, E. V. (2010). *Low income and early cognitive development in the UK: A report for the Sutton Trust*. London: Sutton Trust.

7 APPENDICES

7.1. Appendix A: Variables used in the analysis

Equivalised annual household income (quintiles)

The income that a household needs to attain a given standard of living will depend on its size and composition. For example, a couple with dependent children will need a higher income than a single person with no children to attain the same material living standards. 'Equivalisation' means adjusting a household's income for size and composition so that we can look at the incomes of all households on a comparable basis.

After equivalisation, the sample was split into five, equally sized groups – or quintiles – according to income distribution. Each group thus contains around 20% of families. (For the regression a separate category was created for cases with missing information.)

Area deprivation (SIMD)

Area deprivation is measured using the Scottish Index of Multiple Deprivation (SIMD) which identifies small area concentrations of multiple deprivation across Scotland. It is based on 37 indicators in the seven individual domains of Current Income, Employment, Health, Education Skills and Training, Geographic Access to Services (including public transport travel times for the first time), Housing and a new Crime Domain. SIMD is presented at data zone level, enabling small pockets of deprivation to be identified. The data zones, which have a median population size of 769, are ranked from most deprived (1) to least deprived (6,505) on the overall SIMD and on each of the individual domains. The result is a comprehensive picture of relative area deprivation across Scotland.

In this report, data zones are grouped into quintiles according to their SIMD score. Quintiles are percentiles which divide a distribution into fifths, i.e., the 20th, 40th, 60th, and 80th percentiles. Those respondents whose postcode falls into the first quintile are said to live in one of the 20% least deprived areas in Scotland. Those whose postcode falls into the fifth quintile are said to live in one of the 20% most deprived areas in Scotland.

Further details on SIMD can be found on the Scottish Government website: http://www.scotland.gov.uk/Topics/Statistics/SIMD/Overview

Highest household level of education

At the first wave of data collection for both cohorts, parents were asked to provide information on the nature and level of any school and post-school qualifications they had obtained. This information was obtained for up to two adults in the household

(the main adult respondent and, where applicable, their partner) and was updated at each subsequent contact. Qualifications were grouped according to their equivalent position on the Scottish Credit and Qualifications Framework which ranges from Access 1 to Doctorate. For the purposes of the analysis carried out in chapter 3, these were further banded to create the following categories:

- Lower Standard Grades and below (incl no formal qualifications)
- Upper Standard Grades and intermediate VQs
- Higher Grades, Upper Level VQs and 'Other'
- Degree

The regression analysis reported in chapter 4 used a measure banded into just two categories:

- Degree
- Below degree (incl. cases with missing information)

The highest qualification was defined for each parent and a household level variable was calculated. In couple families this corresponds to the highest qualification among the respondent and his/her partner.

Urban/rural classification

The Scottish Government Urban Rural Classification was first released in 2000 and is consistent with the Government's core definition of rurality which defines settlements of 3,000 or less people to be rural. It also classifies areas as remote based on drive times from settlements of 10,000 or more people. The definitions of urban and rural areas underlying the classification are unchanged.

The classification has been designed to be simple and easy to understand and apply. It distinguishes between urban, rural and remote areas within Scotland and includes the following categories:

- 'Large Urban Areas': Settlements of 125,000 people or more
- 'Other Urban Areas': Settlements of 10,000 to 124,999 people
- 'Accessible Small Towns': Settlements of between 3,000 and 9,999 people and within 30 minutes' drive of a settlement of 10,000 or more
- 'Remote Small Towns': Settlements of between 3,000 and 9,999 people and with a drive time of over 30 minutes to a settlement of 10,000 or more
- 'Accessible Rural': Settlements of less than 3,000 people and within 30 minutes' drive of a settlement of 10,000 or more
- Remote Rural': Settlements of less than 3,000 people and with a drive time of over 30 minutes to a settlement of 10,000 or more

For further details on the classification see the Scottish Government's website:

http://www.gov.scot/Topics/Statistics/About/Methodology/UrbanRuralClassification?utm_source=website&utm_medium=navigation&utm_campaign=statisticsevaluationtools

For the purposes of this report, the above were banded into two categories:

- Urban (large and other urban areas, incl. cases with missing information)
- Small town and rural (accessible small towns, accessible rural, remote small towns, remote rural).

Languages spoken in household

Whether other languages than English were spoken in the household at the time of the sweep 5 (age 5) interview. (Cases with missing information were added to the 'English only' category.)

Parental literacy

At sweep 4 (at the time the cohort child was aged just under 4), the child's main carer was asked two questions designed to measure difficulties with reading and writing. At each question they were asked to indicate whether they had any difficulties with specific tasks. For example, in relation to reading these included understanding what is written in a newspaper and reading aloud from a children's storybook; for writing they included spelling words correctly and making handwriting easy to read.

Responses across all items were combined into a single binary variable indicating whether the child's main carer had any literacy issues. (Cases with missing information were added to the 'No literacy issues' category.)

Parent mental wellbeing

The main carer's mental wellbeing was measured using the Short-Form-12 scale which comprises a physical and a mental wellbeing scale. Data on this measure were collected when the child was aged just under 5 (sweep 5) and when they were in Primary 6 (sweep 8). At each time point, a standardised score was derived which identified those with below average mental wellbeing.

A combined measure then identified those who had below average mental wellbeing at either one or both of the two time points. (Cases with missing information were added to the 'Average or above average' category.)

Parent limiting health problem

Whether the cohort child's main carer reported a limiting long-term health problem at either of the two time points considered in the analysis – i.e. when they were just under 5 (sweep 5) and when they were in Primary 6 (sweep 8). (Cases with missing information were added to the 'No limiting health problem' category.)

School change

Whether the cohort child changed school between the time they started school and the time of their GUS sweep 8 (Primary 6) interview. (Cases with missing information were added to the 'Did not change school' category.)

Change in family type

Whether there was a difference in family type at the two time points considered – when child was aged just under 5 and when they were in Primary 6.

- 'Stable family type': couple or single parent household at both sweeps (incl. cases with missing information)
- 'Change in family type': couple household at one sweep; single parent household at the other.

Significant event happened

Whether the child experienced any of the following events between sweep 5 and sweep 8:

- Death of parent or sibling
- Parent in prison
- Child spent time in care
- Parent lost job

(Cases with missing information were added to the 'No significant changes' category.)

Child limiting health problem

Whether the child was reported by their main carer as having a limiting long-term health problem at either of the two time points considered in the analysis – i.e. when they were just under 5 (sweep 5) and when they were in Primary 6 (sweep 8). (Cases with missing information were added to the 'No limiting health problems' category.)

Child level of social, emotional, behavioural difficulties

On GUS, measures of social, emotional and behavioural development are routinely obtained using items from the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997). A parent report version of the SDQ was included in the self-completion section of the age 5 interview.

The SDQ is a commonly used behavioural screening questionnaire designed for use with children aged between 3 and 16. It consists of 25 questions about a child's behaviour, to which the respondent can answer 'not true', 'somewhat true' or 'certainly true'. Responses can be combined to form five different measures of the child's development, namely emotional symptoms (e.g. excessive worrying), conduct

problems (e.g. often fighting with other children), hyperactivity/inattention (for example, constantly fidgeting), peer relationship problems (e.g. not having close friends) and pro-social behaviour (e.g. being kind to others). Furthermore, the first four measures can be combined into a 'total difficulties' scale.

In this report, a measure of the total difficulties score is used. It was banded using recommended cut-off points. Previously, SDQ scores were most commonly divided into 'normal', 'borderline' and 'abnormal' scores. These bandings were reviewed in 2016 and it is now recommended that SDQ scores on each of the scales are divided into the following categories: 'close to average', 'slightly raised', 'high' and 'very high', with 'very high' indicating multiple problems identified.

The measure used in the report further banded these into two categories:

- Average levels of difficulties ('close to average') (incl. cases with missing information)
- Above average levels of difficulties ('slightly raised', 'high' and 'very high').

Warmth of parent-child relationship

Seven items from the Pianta parent-child relationship scale[17] were used to create a composite measure of the warmth of the parent-child relationship. The child's main carer was asked each item below in the self-completion section of the sweep 7 interview, undertaken just before the child's eighth birthday. For each item the answer options were 'definitely does not apply', 'not really', 'neutral', 'applies sometimes', 'definitely does not apply'.

The following items were used to create the score:

- I share an affectionate, warm relationship with [child]
- [Child] will seek comfort from me
- [Child] values his/her relationship with me
- When I praise [child] he/she beams with pride
- [Child] spontaneously shares information about him/herself
- It is easy to be in tune with what [child] is feeling
- [Child] openly shares his/her feelings and experiences with me

Reliability analysis showed a good internal consistency (alpha = 0.706).

The score was then divided into two categories which were used in the analysis:

- Higher level of warmth

[17] Pianta RC. (1992) *Child–Parent Relationship Scale*. Charlottesville, VA: University of Virginia.

- Lower level of warmth (incl. cases with missing information).

Parent-school interactions

As part of the sweep 7 interview, when the cohort child was just under 8 years old, the main carer was asked if they had attended one or more of the activities listed below in the last approximately two years since their last GUS interview:

- Attending parent evening
- Visiting child's classroom
- Volunteering in school classroom, library, office
- Volunteering for school trip or event
- Offering to volunteer but not asked
- Attending school event where child participated
- Attending school event where child did not participate
- Attending parent council, PTA or school board meeting
- Attending open meeting
- Helping with fundraising

A score was created by adding up the number of activities the parent/carer had attended, which was then banded into the following categories:

- High (7-10 activities)
- Low-Medium (6 activities or less) (incl. cases with missing information).

How often parent helps child look for school-related information

As part of the sweep 7 interview, the child's main carer was asked 'How often do you help [child] look for information about what ^he is learning at school, for example at the library or on the internet?', with answer options 'most days', 'at least once a week', 'a few times a month', 'about once a month', 'a few times a year', 'less often than a few times a year' and 'never'.

The measure used in the analysis had two categories:

- Most days
- Less often (incl. cases with missing information)

Home reading

As part of the sweep 7 interview, the main carer was also asked 'How many days in the last week has [child] looked at books or read stories at home?'

Answers were banded into two categories:

- 6-7 days
- 5 days or less (incl. cases with missing information).

Parent belief in ability to influence child's achievements at school

At sweep 7 the main carer was also asked to what extent they agreed or disagreed with the following statement: 'I believe I can positively influence my child's achievement at school'. Answer categories went from 'Agree strongly' to 'disagree strongly' on a five-point scale. For analysis purposes, the question was coded into two categories:

- Highly positive (agree strongly)
- Less positive (agree, neither agree nor disagree, disagree/disagree strongly, and cases with missing information).

Child's feelings about school

As part of the sweep 7 interview, the child answered a small number of questions on the survey interviewer's laptop. These included three items related to how they felt about school:

- I look forward to going to school
- I hate school
- I enjoy learning at school

Each item had answer options 'never', 'sometimes', 'often', 'always'. The three items were combined into a composite score (internal consistency was good, alpha=0.783) which was then divided into two categories:

- Highly positive (most positive score on all three items)
- Less positive (incl. cases with missing information).

Linked school data

Chapter 4 uses administrative data about the child's Primary 1 school. Consent to link the children's GUS survey data to administrative data - held by the Scottish Government - was obtained from the child's parent or guardian at the 6th sweep of face-to-face data collection, when the child was aged just under 6 years (in 2010/11). Parents/guardians who did not consent at sweep 6, or those who missed an interview at sweep 6, were asked for consent at sweep 7. Consent was captured on a written consent form.

3534 parents (out of 3657, 97%) gave permission to link their survey data with education administrative data at sweep 6. A further 100 (out of 157, 64%) consented at sweep 7. Overall, 3634 of 3814 asked gave consent – 95%. Out of those who consented, 3365 (95%) were successfully matched to education records held by ScotXEd.

After providing the data from GUS, cases were matched by ScotXEd. Matching was done on a sequential basis using all available data and matching to both the 2009 and 2010 pupil census datasets.

Two separate datasets are available: a pupil-level and a school-level dataset. Both datasets contain information relevant to when the GUS children were in Primary 1.

Because children in BC1 straddle two school years, the data in the Primary 1 datasets were not all obtained in the same year – for around three-quarters of children data were obtained for the 2009/10 school year. For the remaining quarter, data were obtained for the 2010/11 school year.

This report uses the following measures:

Size of school: count of the number of pupils enrolled in the school banded as follows:

- 0-100
- 101-200
- 201-300
- 301-400
- Over 400
- [No information].

Proportion of children registered for free school meals: the proportion of pupils in the school registered for free school meals. For the analysis, this measure was split into two categories:

- More than 25%
- 25% or less

In addition, a separate category was created for cases with no information.

Whether denominational school: whether the school was registered as a denominational (faith) school. (Cases with no information were added to the 'No denomination' category).

7.2. Appendix B: Additional descriptive analysis results

Table B-1 Standardised expressive vocabulary score (z score) at Primary 6 - by household income

	Equivalised household income (quintiles)					
	Lowest quintile	**2nd**	**3rd**	**4th**	**Highest quintile**	**All**
Maximum	3.20	3.03	2.67	2.78	3.17	3.24
Bottom 75% of scores	0.48	0.58	0.81	0.86	0.86	0.70
Median	-0.25	0.01	0.15	0.18	0.18	0.05
Bottom 25% of scores	-0.96	-0.78	-0.46	-0.46	-0.42	-0.68
Minimum	-3.19	-3.01	-3.62	-3.19	-2.22	-3.62
Weighted bases	*670*	*536*	*496*	*409*	*407*	*2698*
Unweighted bases	*452*	*498*	*566*	*505*	*532*	*2726*

Base: All children with a valid expressive vocabulary score at sweep 8 / Primary 6 interview with a longitudinal survey weight value. Income details were not provided for 173 cases; therefore, base sizes across the five subgroups do not add up to the total.

Table B-2 Standardised expressive vocabulary score (z score) at Primary 6 - by area deprivation (SIMD)

	Area deprivation (SIMD quintiles)					
	Most deprived quintile	**2nd**	**3rd**	**4th**	**Least deprived quintile**	**All**
Maximum	2.64	2.71	3.17	3.21	3.24	3.24
Bottom 75% of scores	0.34	0.74	0.81	0.87	0.77	0.70
Median	-0.29	0.01	0.18	0.11	0.11	0.05
Bottom 25% of scores	-0.92	-0.75	-0.42	-0.53	-0.56	-0.68
Minimum	-3.19	-2.75	-3.62	-3.19	-2.88	-3.62
Weighted bases	*531*	*493*	*532*	*580*	*563*	*2698*
Unweighted bases	*335*	*430*	*572*	*676*	*713*	*2726*

Base: All children with a valid expressive vocabulary score at sweep 8 / Primary 6 interview with longitudinal survey weight.

Table B-3 Standardised expressive vocabulary score (z score) at Primary 6 - by parental education

	Highest household level of education				
	No quals, Lower std grades and vocational and other quals	**Upper std grades and Intermediate voc quals**	**Higher std grades and Upper level vocational quals**	**Degree level quals**	**All**
Maximum	2.47	3.17	3.21	3.24	3.24
Bottom 75% of scores	0.08	0.60	0.62	0.91	0.70
Median	-0.49	-0.18	0.05	0.25	0.05
Bottom 25% of scores	-1.29	-0.85	-0.72	-0.35	-0.68
Minimum	-3.19	-3.62	-2.94	-2.94	-3.62
Weighted bases	*257*	*485*	*898*	*1054*	*2698*
Unweighted bases	*153*	*387*	*886*	*1296*	*2726*

Base: All children with a valid expressive vocabulary score at sweep 8 / Primary 6 interview with longitudinal survey weight. Parental education details were not available for 4 cases; therefore, base sizes across the four subgroups do not add up to the total.

7.3. Appendix C: Multivariable analysis results

Description of the analysis undertaken

Linear regression analysis

Many of the factors we are interested in are related to each other as well as being related to cognitive ability. For example, parents on lower incomes are also more likely to have lower levels of education and to live in areas of high deprivation. Simple analysis may identify a relationship between income and language ability. However, this relationship may be occurring because of the underlying association between income and education. Thus, it may be the lower level of education among lower-income parents which is associated with a greater likelihood of lower language ability in their children rather than the fact that they are poor. To avoid this difficulty, multivariable regression analysis was used. This analysis allows the examination of the relationships between a dependent (outcome) variable and multiple independent (explanatory) variables whilst controlling for the inter-relationships between each of the independent variables. This means it is possible to identify an *independent* relationship between any single explanatory variable and the outcome variable; to show, for example, that there is a relationship between income and language ability that does not simply occur because parental education and income are related.

The regression models developed for this report were fitted with standardised WIAT-II vocabulary score (z score) measured when the child was in Primary 6 as the dependent variable. Standardised BAS-II vocabulary score (z score) measured at age 5 was included as an independent variable. Measures of social background characteristics and demographics, and various additional factors identified from the literature were also added as independent variables. By including a measure of ability at age 5, the results of this analysis identify characteristics which are associated with a relative change in assessment score between age 5 and Primary 6, after controlling for other, potentially confounding, characteristics. Note, though, that the identification of associations between one or more independent variables and a dependent variable does not necessarily imply that the independent variable(s) *causes* the dependent variable (the outcome).

The characteristics, experiences and circumstances considered in the analysis are outlined in Table 4-1. Readers should note that to ensure consistency in the analysis, for variables with a high number of cases with missing values (e.g. income), a separate category ('No information') was created. For cases with smaller numbers of missing cases (~<100), cases with missing values were added to the modal (most common) category. Further details are provided in Appendix A.

Note also that only children with valid vocabulary scores at both time points were included in the analysis (36 children with a valid vocabulary score at age 5/Primary 1 were excluded from the analysis because there was no valid vocabulary score at Primary 6). Furthermore, data were weighted using the GUS longitudinal survey

weight, meaning that only cases which have taken part in every face-to-face sweep of GUS up to and including sweep 8 were included. In total, 2726 children were included in the analysis.

The regression analysis was carried out in four stages:

- **Stage 1: Univariate linear regression models** (Tables C-1 to C-24)
    - To examine the relationship between the two standardised vocabulary scores used in the analysis, first, a univariate linear regression model was fitted with standardised expressive vocabulary score at Primary 6 as the dependent variable and standardised expressive vocabulary score at Primary 1 as the only independent variable.
    - Individual linear regression models were then fitted for each of the factors outlined in Table A. In each of these models standardised expressive vocabulary score at Primary 6 was the dependent variable, and standardised expressive vocabulary score at Primary 1 was included as a covariate.
- **Stage 2: Multivariable model with Stage 1 significant factors** (Table C-25)
    - The next stage of analysis involved entering the factors which were significant at the 90% level into a single regression model. In so doing, this analysis explored the extent to which each factor remained independently associated with a relative improvement or decline in language ability over the primary school period once controlling for the influence of other factors, including social background.
- **Stage 3: Multivariable model with Stage 2 significant factors** (Table C-26)
    - In the third stage of the analysis, a final model was created including only those factors which were significant at the 90% level in the stage 2 model. This is referred to as the 'final model'.
- **Stage 4: Stage 3 multivariable model with interaction effects** (Table C-27)
    - To explore whether associations differed according to parental education, interaction effects were fitted to the stage 3 model (the 'final model') between parental education and each of the independent variables except for Primary 1 vocabulary score.

Interpreting the tables

The weighted sample size for each category is provided in the 'Weighted base' column. The sample size given in the top row for each variable is the sample size for the reference category, which is given in brackets.

All figures quoted in this report have a margin of error because they are estimates based on a sample of children, rather than all children. The p-value is an estimation

of how likely it is that we would find a relationship in our sample of children if there was no actual relationship in the population (i.e., broadly speaking, among children in Scotland who are the same age as the GUS children but who are not part of GUS). Thus, the smaller the p-value ($p<0.05$), the more confident we can be that our results are likely to apply to children in Scotland more widely.

For continuous independent variables (covariates) the regression coefficient ('Coeff') illustrates the relative level of change (positive or negative) in language ability score at P6 if score at P1 is increased by 1 unit. A significant ($p<0.05$) positive coefficient denotes a relative improvement in ability score and a significant negative coefficient denotes a relative decline in ability score for every one-unit increase in P1 score.

For categorical independent variables (factors) the regression coefficient ('Coeff') illustrates the relative level of difference (positive or negative) in language ability for each sub-group as compared with the reference sub-group. A significant ($p<0.05$) positive coefficient denotes a higher ability score and a significant negative coefficient denotes a lower ability score when compared with the reference sub-group. The reference sub-group is indicated in brackets.

The 95% confidence interval is an indication of the level of uncertainty in the coefficient estimate.

Table C-1 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1

	Weighted base	p-value	Coeff	95% Confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	0.431	0.385	0.476
R squared		0.172			
Total N (unweighted)	*2726*				
Total N (weighted)	*2698*				

Table C-2 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and child's gender

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.432	.386	.478
Child's gender					
Male	*1374*	.365	.035	-.042	.112
Female (ref)	*1324*	-	-	-	-
R Squared		0.172			
Total N (unweighted)	*2726*				
Total N (weighted)	*2698*				

Table C-3 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and equivalised annual household income

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.411	.363	.459
Equivalised annual household Income		.025			
No Information	*160*	.803	.027	-.189	.243
Top Quintile (>=£37,857)	*358*	.001	.244	.099	.389
4th Quintile (>=£29,126<£37,857)	*525*	.041	.150	.007	.293
3rd Quintile (>=£19,643<£29,126)	*430*	.035	.175	.012	.337
2nd Quintile (>=£12,217<£19,643)	*605*	.132	.097	-.030	.225
Lowest Quintile (<£12,217) (ref)	*620*	-	-	-	-
R Squared		0.178			
Total N (unweighted)	*2726*				
Total N (weighted)	*2698*				

Table C-4 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and level of area deprivation (SIMD)

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.419	.373	.466
Area Deprivation (SIMD)		.049			
Least Deprived Quintile	*523*	.029	.153	.016	.290
4th Quintile	*556*	.026	.153	.019	.286
3rd Quintile	*510*	.002	.222	.085	.358
2nd Quintile	*508*	.078	.122	-.014	.259
Most Deprived Quintile (ref)	*602*	-	-	-	-
R Squared		0.178			
Total N (unweighted)	*2726*				
Total N (weighted)	*2698*				

Table C-5 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and parental education

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.411	.364	.458
Highest level of parental education					
Degree level or above	*964*	.000	.180	.100	.260
Below degree (incl. missing) (ref)	*1734*	-	-	-	-
R Squared		0.179			
Total N (unweighted)	*2726*				
Total N (weighted)	*2698*				

Table C-6 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and urban/small town or rural location

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.424	.379	.469
Urban/rural location					
Small town or rural	*852*	.005	.139	.045	.233
Urban (incl. missing) (ref)	*1846*	-	-	-	-
R Squared		0.176			
Total N (unweighted)	*2726*				
Total N (weighted)	*2698*				

Table C-7 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and languages spoken in the household

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.434	.387	.481
Languages spoken in household					
Other language(s) spoken	*141*	.120	.133	-.036	.302
English only (incl. missing) (ref)	*2558*	-	-	-	-
R Squared		0.173			
Total N (unweighted)	*2726*				
Total N (weighted)	*2698*				

Table C-8 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and parent literacy issues

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.430	.384	.477
Parent literacy issues					
One or more literacy issues	*338*	.866	-.013	-.171	.145
No literacy issues (incl. missing) (ref)	*2361*	-	-	-	-
R Squared		0.172			
Total N (unweighted)	*2726*				
Total N (weighted)	*2698*				

Table C-9 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and parent mental wellbeing

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.433	.386	.480
Parent mental wellbeing					
Below average at sweep 5 and/or sweep 8	*632*	.456	.036	-.060	.132
Average or above at both sweeps (incl. missing) (ref)	*2066*	-	-	-	-
R Squared		0.172			
Total N (unweighted)	*2726*				
Total N (weighted)	*2698*				

Table C-10 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and whether parent has limiting health problem

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.431	.385	.478
Parent limiting long-term health problem					
Parent had limiting health problem at sweep 5 and/or sweep 8	*356*	.370	.056	-.068	.179
Parent had no limiting health problem (incl. missing) (ref)	*2342*	-	-	-	-
R Squared		0.172			
Total N (unweighted)	*2726*				
Total N (weighted)	*2698*				

Table C-11 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and whether child has a limiting health problem

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.430	.384	.477
Child limiting long-term health problem					
Child had limiting health problem at sweep 5 and/or sweep 8	*269*	.937	-.007	-.174	.161
Child had no limiting health problem (incl. missing) (ref)	*2429*	-	-	-	-
R Squared		0.172			
Total N (unweighted)	*2726*				
Total N (weighted)	*2698*				

Table C-12 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and whether child had above average levels of social, emotional and behavioural difficulties

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.418	.371	.465
Child's social, emotional & behavioural difficulties					
Above average difficulties at sweep 5 and/or sweep 8	*433*	.001	-.198	-.310	-.087
Average levels of difficulties at both sweeps (incl. missing) (ref)	*2264*	-	-	-	-
R Squared		0.177			
Total N (unweighted)	*2726*				
Total N (weighted)	*2698*				

Table C-13 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and whether child experienced parental separation or re-partnering

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.430	.384	.475
Parental separation/re-partnering					
Change in family type	*432*	.600	-.036	-.171	.100
Stable family type (incl. missing) (ref)	*2266*	-	-	-	-
R Squared		0.172			
Total N (unweighted)	*2726*				
Total N (weighted)	*2698*				

Table C-14 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and whether child changed school

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.429	.383	.475
Whether child changed school					
Changed school	*419*	.279	-.063	-.180	.053
Did not change school (incl. did not attend school and missing) (ref)	*2280*	-	-	-	-
R Squared		0.173			
Total N (unweighted)	*2726*				
Total N (weighted)	*2698*				

Table C-15 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and whether child experienced significant adverse life event

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.431	.385	.476
Whether child experienced significant adverse life event					
Significant adverse event occurred	*327*	.818	.015	-.116	.147
No significant adverse event (incl. missing) (ref)	*2371*	-	-	-	-
R Squared		0.172			
Total N (unweighted)	*2698*				
Total N (weighted)	*2726*				

Table C-16 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and child's feelings about school

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.431	.385	.477
Child's feelings about school					
Highly positive	*824*	.335	.042	-.044	.128
Less positive (incl. missing) (ref)	*1874*	-	-	-	-
R Squared		0.172			
Total N (unweighted)	*2698*				
Total N (weighted)	*2726*				

Table C-17 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and size of P1 school

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	2698	.000	.431	.384	.478
Number of pupils in P1 School		.228			
No information	*205*	.445	.075	-.121	.271
0-100	*215*	.538	.057	-.127	.240
101-200	*629*	.495	.058	-.111	.227
201-300	*678*	.619	-.043	-.214	.129
301-400	*688*	.390	-.065	-.215	.085
More than 400 (ref)	*284*	-	-	-	-
R Squared		0.175			
Total N (unweighted)	*2698*				
Total N (weighted)	*2726*				

Table C-18 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and whether P1 school denominational

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	2698	.000	.431	.385	.477
Whether P1 school a denominational school					
Non-denominational (incl. no information)	*2011*	.613	-.024	-.121	.072
Denominational (any religion) (ref)	*687*	-	-	-	-
R Squared		0.172			
Total N (unweighted)	*2698*				
Total N (weighted)	*2726*				

Table C-19 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and % of children at P1 school registered for free school meals

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	2698	.000	.422	.375	.470
% of pupils in P1 school registered for FSM		.095			
No information	*205*	.464	.057	-.098	.213
More than 25%	*633*	.068	-.113	-.236	.009
25% or less (ref)	*1860*	-	-	-	-
R Squared		0.175			
Total N (unweighted)	*2698*				
Total N (weighted)	*2726*				

Table C-20 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and warmth of parent-child relationship

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	2698	.000	.430	.384	.476
Parent-child warmth					
High level of warmth	*916*	.440	.028	-.044	.101
Lower level of warmth (incl. missing) (ref)	*1782*	-	-	-	-
R Squared		0.172			
Total N (unweighted)	*2698*				
Total N (weighted)	*2726*				

Table C-21 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and parental interactions with child's school

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	2698	.000	.425	.379	.470
Parent interactions with child's school					
High (7-10 interactions)	*510*	.005	.112	.036	.189
Low-Medium (0-6 interactions) (incl. missing) (ref)	*2189*	-	-	-	-
R Squared		0.174			
Total N (unweighted)	*2698*				
Total N (weighted)	*2726*				

Table C-22 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and how often parent helps child look for school-related information

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	2698	.000	.430	.384	.476
How often parent helps child look for school-related information					
Most days	*493*	.333	-.059	-.180	.062
Less often (incl. missing)	*2205*	-	-	-	-
R Squared		.173			
Total N (unweighted)	*2698*				
Total N (weighted)	*2726*				

Table C-23 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and home reading

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.420	.372	.468
Home reading in last week					
Most days (6-7 days) (incl. missing)	*1557*	.002	.139	.053	.225
5 days or less (ref)	*1141*	-	-	-	-
R Squared		.177			
Total N (unweighted)	*2698*				
Total N (weighted)	*2726*				

Table C-24 Linear regression model predicting standardised expressive vocabulary score at P6, by standardised expressive vocabulary score at P1 and parental belief they can influence child's achievements at school

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	2698	.000	.427	.381	.473
Parent belief they can influence child's achievements at school					
Highly positive (strongly agree)	*1116*	.089	.066	-.010	.143
Less positive (incl. missing) (ref)	*1582*	-	-	-	-
R Squared		.173			
Total N (unweighted)	*2698*				
Total N (weighted)	*2726*				

Table C-25 Linear regression model predicting standardised expressive vocabulary score at P6 - by factors individually associated with change in univariate analysis

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.384	.335	.433
Child's gender					
Boy	*1374*	.122	.060	-.017	.137
Girl (ref)	*1324*	-	-	-	-
Area deprivation (SIMD)		.224			
Least deprived quintile	*523*	.884	-.012	-.175	.151
4th quintile	*556*	.938	-.006	-.162	.150
3rd quintile	*510*	.157	.103	-.041	.248
2nd quintile	*508*	.445	.052	-.084	.189
Most deprived quintile (ref)	*602*	-	-	-	-
Equivalised annual household Income		.555			
No information	*160*	.748	-.034	-.242	.174
Top quintile (>=£37,857)	*358*	.126	.128	-.037	.293
4th quintile (>=£29,126<£37,857)	*525*	.591	.042	-.112	.195
3rd quintile (>=£19,643<£29,126)	*430*	.188	.103	-.052	.258
2nd quintile (>=£12,217<£19,643)	*605*	.435	.052	-.080	.184
Lowest quintile (<£12,217) (ref)	*620*	-	-	-	-
Highest level of parental education					
Degree level or above	*964*	.012	.118	.026	.209
Below degree (incl. missing) (ref)	*1734*	-	-	-	-
Urban/small town or rural location					
Small town or rural	*852*	.021	.118	.018	.218
Urban (incl. missing) (ref)	*1846*	-	-	-	-
% of pupils in P1 school registered for free school meals		.799			
No information	*205*	.508	.051	-.101	.202
More than 25%	*633*	.908	.008	-.130	.146
25% or less (ref)	*1860*	-	-	-	-
Child's social, emotional and behavioural difficulties					
Above average difficulties at sweep 5 and/or sweep 8	*433*	.011	-.163	-.286	-.039

Table C-25 continued	Weighted base	p-value	Coeff	95% confidence interval	
Average levels of difficulties at both sweeps (incl. missing) (ref)	*2265*	-	-	-	-
Home reading in last week					
Most days (6-7 days) (incl. missing)	*1557*	.023	.107	.015	.200
5 days or less (ref)	*1141*	-	-	-	-
Parent interactions with child's school					
High (7-10 interactions)	*510*	.340	.035	-.038	.109
Low-Medium (0-6 interactions) (incl. missing) (ref)	*2189*	-	-	-	-
Parent belief they can influence child's achievements at school					
Highly positive (strongly agree)	*1116*	.453	.029	-.047	.105
Less positive (incl. missing) (ref)	*1582*	-	-	-	-
R Squared		0.195			
Total N (unweighted)	*2726*				
Total N (weighted)	*2698*				

Table C-26 Linear regression model predicting standardised expressive vocabulary score at P6 – final model

	Weighted base	p-value	Coeff	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.391	.342	.440
Child's gender					
Boy	*1374*	.119	.063	-.017	.142
Girl (ref)	*1324*	-	-	-	-
Highest level of parental education					
Degree level or above	*964*	.001	.143	.061	.225
Below degree (incl. missing) (ref)	*1734*	-	-	-	-
Urban/small town or rural location					
Small town or rural	*852*	.013	.124	.027	.220
Urban (incl. missing) (ref)	*1846*	-	-	-	-
Child social, emotional & behavioural difficulties					
Above average difficulties at sweep 5 and/or sweep 8	*433*	.002	-.179	-.290	-.068
Average levels of difficulties at both sweeps (incl. missing) (ref)	*2265*	-	-	-	-
Home reading in last week					
Most days (6-7 days) (incl. missing)	*1557*	.012	.118	.027	.209
5 days or less (ref)	*1141*	-	-	-	-
R Squared		.191			
Total N (unweighted)	*2726*				
Total N (weighted)	*2698*				

Table C-27 Linear regression model predicting standardised expressive vocabulary score at P6 – final model with interaction effects

	Weighted base	p-value	Regression coefficient	95% confidence interval	
Standardised vocabulary score (P1)	*2698*	.000	.391	.342	.440
Child's gender		.060			
Boy	*1374*	.713	.019	-.085	.124
Girl (ref)	*1324*	-	-	-	-
Highest level of parental education		.138			
Degree level or above	*964*	.566	.047	-.116	.210
Below degree (incl. missing) (ref)	*1734*	-	-	-	-
Urban/small town or rural location		.017			
Small town or rural	*852*	.015	.153	.031	.275
Urban (incl. missing) (ref)	*1846*	-	-	-	-
Child social, emotional & behavioural difficulties		.004			
Above average difficulties at sweep 5 and/or sweep 8	*433*	.021	-.167	-.307	-.026
Average levels of difficulties at both sweeps (incl. missing) (ref)	*2265*	-	-	-	-
Home reading in last week		.003			
Most days (6-7 days) (incl. missing)	*1557*	.208	.079	-.045	.204
5 days or less (ref)	*1141*	-	-	-	-
Interaction effects					
Parental education* Child's gender	-	.156	-	-	-
Parental education * Urban/small town or rural location	-	.376	-	-	-
Parental education * Child social, emotional and behavioural difficulties	-	.850	-	-	-
Parental education * How often parent reads with child	-	.223	-	-	-
R Squared		.192			
Total N (unweighted)	*2726*				
Total N (weighted)	*2698*				

ScotCen
Social Research that works for society